THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR ....Lorraine Joan MacGregor....

TITLE OF THESIS ....Aggregation and Transferability of

Mode Choice Models in Urban Space:

The Case for Ottawa....

DEGREE FOR WHICH THESIS WAS PRESENTED ....Master of Arts....

YEAR THIS DEGREE GRANTED ....1981....

THE UNIVERSITY OF ALBERTA

AGGREGATION AND TRANSFERABILITY OF MODE CHOICE MODELS IN URBAN SPACE: THE CASE FOR OTTAWA

by

© Lorraine Joan MacGregor

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF ARTS

DEPARTMENT OF ECONOMICS

EDMONTON, ALBERTA

SPRING 1981

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Transferability and Aggregation of Mode Choice Models in Urban Space: The Case for Ottawa," submitted by Lorraine Joan MacGregor, in partial fulfilment of the requirements for the degree of Master of Arts in Economics.

This thesis is dedicated to Gordon and Margaret, and especially to my mother and father, whose love and constant support has given each of us the confidence to take advantage of our opportunities and to live life to the fullest.

ABSTRACT

A priori, one would expect individual responsiveness to policy variables to vary according to location in the urban area. The intent of this thesis is to examine how population elasticities vary across urban space, and the reasons for this variation. This expected variation could be due to differences between choice environments facing individuals, differences in observed preferences of those individuals, or both.

Any analysis to this end must involve a choice theory and a corresponding aggregation procedure which establishes a *causal* link from the characteristics of choice settings and individuals' characteristics to specific observed aggregate behavior. Summaries of aggregate response, such as elasticities, must therefore be sensitive to these factors. A disaggregate Logit Model and a corresponding aggregation procedure developed by Westin [60] can achieve this result, although due to new sampling requirements, these procedures have not yet been widely adopted in city planning procedures in Canada.

Disaggregate choice theory underlies the Logit Model which is used in this analysis. An individual choice probability can be derived, which has a structure which links that probability to characteristics of the individual and his choice environment.

Observed aggregate behavior is the sum of the choices of the individuals in the population. Westin's aggregation procedure preserves this causality, and gives the further advantage of allowing one to

separate the model of individual choice from the aggregation procedure. The parameters of the model of individual choice are seen to be parameters of observed preferences, which under certain conditions may be similar across individuals. The aggregation procedure accounts for differences in choice settings and translates these differences to differences in aggregate demand.

If the parameters of observed preferences can be shown to be stable across populations, then potential exists for borrowing these results between populations. This is known as model transferability, and if successful provides significant information economies to policy makers.

Theoretical and practical preconditions for successful transferability exist, and must be known to the policy maker before models can be transferred. If these conditions hold, then predictions can be made about new areas using information borrowed from the models of individual choice, and combined with new area specific data in the aggregation procedure for use in forecasting and elasticities for policy predictions.

ACKNOWLEDGEMENTS

I would like to thank Dr. David Gillen for his patience and the many valuable insights that he has so generously shared. I would further like to thank Dr. Soren Lemche for his helpful comments, and Art for his help and his very special support and friendship. I would especially like to thank David, Soren and Art for their own examples in the pursuit of excellence in what they do--which has been invaluable to me.

I would also like to acknowledge the patient and generous assistance of Alon Sharpe on the computer and Mrs. Doreen Nordquist for her obvious typing talents.

Finally, I must express my thanks to Mr. Michael Fellows, whose contributions towards this thesis are, I think, known to him, to the Brand family for their generosity, and to the Transportation Development Agency for their interest in this research and their generous financial assistance for my graduate education.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

TABLE OF CONTENTS (CONT'D)

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$A = (a = 1,2,\ldots,A)$	index set of individuals
$I = (i = 1,2,\ldots,I)$	index set of groups of individuals
A^i	the ith group of individuals
a^i	individual in A^i
$s'[a]$	vector of all socioeconomic characteristics of an individual
$s[a^i]$	vector of observable socioeconomic characteristics of individual a^i
$s'[a^i]$	vector of unobservable socioeconomic characteristics of individual a^i
$J = (j = 1,2,\ldots,J)$	index set of alternatives
x_j	vector of observable characteristics of the jth alternative
x'_j	vector of unobservable characteristics of the jth alternative
$K = (k = 1,2,\ldots,K)$	index set of characteristics

Chapter 1

INTRODUCTION

This thesis investigates and analyzes individual behavior within an urban economy, as these individuals engage in a specific type of economic activity. The analysis concerns the establishment of a link between the structure of the individual decision-making processes and specific observed behavior--both for the purposes of description and ultimately to evaluate the effects of alternative public policies on individual choices.

1.1 OUTLINE OF THE MAIN ISSUES UNDERLYING THE ANALYSIS

1.1.1 Traditional Analysis and the Notion of Continuity

Economic analysis of individuals necessarily includes a number of assumptions regarding their willingness and ability to make choices.

1. Individuals are aware of their needs.
2. Individuals have a consistent motivation in the satisfaction of their wants--the maximization of utility. This implies that individuals will be faced with the problem of optimizing within their environmental constraints.
3. Individuals have knowledge of options available for the satisfaction of a particular set of wants, and they are able to compare competing alternatives.
4. Individuals are rational. They behave in accordance with a set of preferences which are fixed in the short run.

5. Different choice situations will explicitly include or exclude the state of uncertainty, either with respect to prices, or regarding the utility from the commodity consumed.

From assumptions such as these, demand analysis usually proceeds to the determination of a specific representation of individual preferences and constraints, ultimately leading to the determination of a set of optimizing rules for the consumer.

The form of the theoretical analysis of consumer demand can follow various perspectives. However, consistent in most theoretical analyses of demand is the notion of continuity--consumers making decisions at the margin. Continuous adjustments in demand are made in reaction to even very small price changes.

As demand analysis is a crucial input into marketing decision-making processes, there has been a transition from, in many cases, very elegant theory to models which can be practically applied to demand-oriented economic problems. The assumption of continuity has been maintained in the transfer of the theory of demand to models intended for practical application. Therefore, considerable effort has been expended in the search for precise forms of demand relationships in various consumer situations. This has implied the need for the enumeration of relevant explanatory variables and an adequate representation of the form or manner in which they affect choices.

The practical problems associated with the empirical identification of demand relationships are numerous and well documented. The nature of these problems has led many to a distrust or impatience with the traditional theoretical notion of the demand function. However, Baumol [3] clearly reminds us that within the context of traditional

marginal analysis, there exists no better substitute.[1]

For many, however, there has been a rejection of these traditional procedures for certain types of analyses. This is in recognition of an important class of decision which the traditional analysis cannot address. These are decisions related to choices between discrete alternatives. This recognized inadequacy in the theory has led to successful attempts to theoretically reanalyze the determinants of these types of choices.

One area of demand modeling which has helped initiate this theoretical concern has been the examination of the household in an urban area, where the application of marginal analysis to some types of behavior is questionable.

For instance, the spatial differentiation of goods and services in an urban area gives rise to a derived demand for transportation. Observed behavior in transportation is the result of a utility maximizing motivation, but many transportation decisions are discrete choices. They should not be modeled by means consistent with only traditional marginal analysis.

Recognition and concern regarding the nature of transportation choices has not always been reflected in transportation and urban modeling procedures. Much of the modeling contained in urban transportation planning procedures reflects an orientation toward the description of observed aggregate flows of traffic between areas, ignoring underlying discrete decisions at the level of the individual. These models have been used extensively to provide demand-oriented information to transportation planners. But ironically their most important inadequacy is in their neglect to address the issues

important to such analyses, such as the determinants and structure of such decisions.

Increased research into the theories of consumer behavior to this end has led to the rise of new forms of choice theory (as opposed to "demand" theory). This theory links the structure of the decision-making process to specific observed behavior, and in doing so respects the role of individuals as determinants of observed aggregate behavior in the economy.

Disaggregate modeling represents a procedure in which the characteristics of individuals and commodities are combined into a proposed decision-making process, which then results in a particular structure of a probability of taking a particular action. The model provides a causal link between the characteristics of the individuals, the structure of the individual choice probabilities and (with appropriate aggregation procedures) observed aggregate market responses.

Demand elasticities are traditionally considered to be summaries of aggregate demand responses. At the aggregate level, all choices, regardless of their structure, appear to be continuous. Traditional elasticities, using the notion of a representative consumer, are applied to aggregate data, and are often applied without distinction being made as to the sources of apparent aggregate adjustment in demand. There exist, however, different scenarios, which appear at an aggregated level to be identical[2]--ranging from situations where all adjustment is continuous marginal adjustment to price changes, to situations where some of the change in consumption in a population is quantity adjustment by individuals, and part is commodity switching by others. At the other extreme, all of the adjustments to a price change could be commodity

switching.

The application of traditional aggregated elasticities indiscriminately to situations where they do not adequately reflect the structure of the aggregate changes will result in a potential loss of accuracy and information about that population.

The disaggregate modeling procedure, by delaying aggregation until after estimation of the model, allows for the redefinition of the elasticity--leaving it more sensitive to differences between individuals and supply side characteristics within different populations. And importantly, the structure of the individual's choice process is inherent in the structure of the elasticity, increasing its consistency with the situation of discrete individual choices underlying observed aggregate flows.

An elasticity which is sensitive to differences between individuals and supply side characteristics can then be applied to a variety of population subgroups without compromising the individuality of those populations. This allows the elasticities to be potentially more sensitive to specific modeling contexts required by different policy questions.

1.1.2 Transferability

In addition to potentially more accurate aggregate analysis for discrete choice problems, another important purpose is served by disaggregate choice models. This is in terms of the information afforded about the preferences and behavior of the individuals themselves.

As stated, traditional marginal analysis and disaggregate choice models differ in their respect of the individuality of each consumer.

Traditional models impose the similarity of consumers into the model structure through the restrictive assumptions about the representative consumer required for certain aggregation procedures. Paradoxically, because of the aggregate nature of these models, the model results are dependent upon the context in which the model is applied, which precludes interpopulation application of results from that model which assumes that individuals are identical. This has practical implications in that the models must be applied separately to different populations, and model results from one cannot be applied to any other than an identical population.

The disaggregate approach differs on two accounts.

1. Only part of the preferences of individuals are theoretically considered to be representative of other consumers. The theory allows a portion of the individual's preferences to be personal and identifiable only to that particular consumer.
2. A distinction can be made between what is *theoretically* assumed regarding representative preferences and what is *identifiable* as representative from the estimated model. This is possible because both the theoretical and the practical analyses take place at the level of the individual.

If any portions of the representative preferences are identifiable (or if the conditions are known under which this might be the case), then this acquired information can be used to represent other populations in forecasts and/or elasticities requiring this information.

By allowing for the individuality of the consumer, a model is derived which potentially yields information which can be used to describe aspects of other individuals or populations. In contrast,

traditional models incorporate the assumption of representativeness but cannot yield information to turn this assumption into practice.

1.2 THE ISSUES TO BE EXAMINED

One intuitively expects different types of behavior from different groups of individuals who live in a heterogeneous urban area. The structure of elasticities derived from disaggregate theory confirms that at least some portion of the population differences will be captured in those aggregate measures.

Conceptually, population can be differentiated on the basis of almost any criteria. For a policy maker it is logical to do so on the basis of planning related variables. It may be expedient in some cases to differentiate populations spatially, so that some insight can be gained into the efficiency of using city-wide policy. A major premise of the thesis, then, is that measures of responsiveness to transportation variables--the aggregate elasticities--will vary as one moves across urban space.

The analysis will attempt to examine why this is so. This will be accomplished through examination of the components of the aggregate elasticities to see which components vary systematically across spatially differentiated populations. For instance, the observable preferences of individuals might be the same, but the elasticities might differ because of differences in the parameters of the distributions of characteristics which are also included in the elasticity measures. However, the elasticities might differ because of differences in both the preferences of individuals as measured and differences in the distributions of characteristics of the population

and the alternatives offered.

If the first of these two situations results, then the model may have identified the representative portion of utility, which is theoretically assumed.[3] This will have desirable implications for policy makers wishing to economize in forecasting between populations, or those wishing to apply the disaggregate procedures where no specific disaggregate data set for the population is available.

If the second case results, then every population would have to be modeled separately because as far as could be identified in the model their preferences would share little in common with other populations. In this case less optimistic results are implied for policy makers trying to economize in prediction methodology.

1.3 OVERVIEW OF THE ANALYSIS

The analysis will be done by:

1) Estimating the logit model for an entire urban area and calculating aggregate elasticities for the area on the basis of these results.
2) The logit model will be estimated separately on each subpopulation, and aggregate elasticities calculated for each one, using parameters of preferences and characteristics distributions from each corresponding model. These will be compared to the elasticity in 1) above.
3) Aggregate elasticities will be calculated for spatially-differentiated subpopulations using parameters of preferences borrowed from the city-wide model and distributional information specific to the population subgroups. These will be compared to each other and the elasticities in 1) and 2) above.

4) New probability distributions for each subarea will be generated, where each individual will be ascribed a hypothetical choice probability generated from coefficients borrowed from other models. These new probabilities will then provide the basis for new elasticities. These generated measures embodying transferred results will be compared to the actual probabilities in each subarea.
5) In addition, a brief comparison will be made as to any differences in transferability of *aggregate results* (i.e., the similarities of the elasticities) as opposed to similarity of the model parameters (coefficients) reflecting the representative preferences. Much of the existing literature regarding transferability uses similarity of the model parameters as the criterion for successful model transferability. However, the use of model results is for aggregate prediction, and there may be some room for reinterpretation of how good "transferability" should be measured.

1.4 CHAPTER SUMMARY

In Chapter 2 of this thesis, the disaggregate model will be theoretically described. The assumptions underlying the theory will be presented, and issues and implications arising from the procedure will be discussed. The specific form of the model derived will be the Logit Model.[4]

In Chapter 3 the implications of the Logit Model will be discussed in terms of aggregate behavior. Some of the issues with respect to the identification of a link from individual behavior to aggregate behavior will be clarified. The aggregation procedure to be used will

be identified, and the structure of the aggregate elasticities which will be used will be derived.

The structure of the aggregate measures to be employed will be seen to include parameters of the potentially identifiable representative portions of utility. This chapter will also include a discussion of the issues related to transferability--the use of model results about one population in the representation of another population. The conditions under which the representative portions of utility can be identified will be discussed. This will provide information towards potential justification for the transfer of information about one population for use in forecasts or elasticities describing another.

The various issues related to the specification of the representative components of utility will be discussed in Chapter 4. This chapter will serve as the reference for the actual selection of variables for the model.

Specifically the Logit Model will be applied to the problem of choice of mode for urban work trips, and will be applied to data collected for the City of Ottawa. Chapter 5 will include discussion of the issues related to the data set and the estimation procedure. This chapter will also contain the empirical application of the model to the choice of mode for urban work trips, with a presentation of model results and elasticities for the City of Ottawa and its subpopulations. Summaries and conclusions will be presented in Chapter 6.

Chapter 1

FOOTNOTES

[1]Baumol [3], p. 227.

[2]This is also a theme expressed by Novshek and Sonnenschein [38] in their discussion of a new structure for aggregate elasticities.

[3]This means that this portion of their utility function is the same, even if their utility level is different.

[4]The type of model will be seen to be defined in terms of a specific derived form of decision probabilities.

Chapter 2

THE THEORY OF INDIVIDUAL CHOICE

Economic theory is essential input to the provision of sound guidelines for solutions to problems associated with policy decisions in the economy. Increased interest in microeconomic problems has placed pressure on the ability of microeconomic modeling to adequately reflect the structure of particular microeconomic decisions. This has led many to question the applicability of traditional demand theory to help solve contemporary microeconomic problems.

The present status of traditional theory is well described by James Heckman:

> As economists attempt to make greater use of their theory to solve such practical problems as estimating the demand for new modes of travel, and ascertaining the determinants of the labor supply of women, the analytical fiction of the representative consumer and its economic analogue, the classical regression model, have been less useful.
>
> Increasingly, economists have come to realize that choices at the extensive margin (discrete choices) are just as interesting and often of greater empirical importance than the analysis of choice at the intensive margin that is treated in the traditional analysis.
>
> Because the source of sampling variations critically affects the formulation of many models of discrete choice, the traditional schizophrenia of Marshallian Econometrics that separates the formulation of an economic model for a typical individual from its stochastic specification is absent from many of the best papers in the literature.[1]

Disaggregate behavioral demand modeling is an attempt to resolve this inadequacy of traditional demand theory to represent the structure of a very important class of choices--those which are discrete (more formally known as "quantal response") and which marginal analysis cannot

explain.[2] These are choices where the consumer makes "all or nothing" decisions on the basis of total willingness to pay for a good, as opposed to decisions relating to how much of a commodity to consume, which traditionally relates to marginal expenditures.[3]

The disaggregate approach, which has been somewhat revolutionary in demand analysis in the direct treatment of this phenomenon, has been successfully applied to such aspects of choice as urban travel demand, college selection, voting behavior, migration, the demand for consumer durables, and occupational choice.[4]

2.1 INTRODUCTION TO THE DISAGGREGATE METHODOLOGY

In this analysis, the demand for new and existing commodities is examined in terms of the utility over jointly-consumed characteristics, rather than in terms of the utility of the good itself.[5] It is hypothesized that if an individual is given an opportunity to choose from a set of characteristics collections, the individual will choose that collection which maximizes his utility. Since the characteristics are offered to the consumer in a finite number of bundles, discontinuous choices between commodities in different choice settings will be observed. It is the role of policy to systematically affect the thresholds of individuals causing substitution between alternatives in order to appreciably affect their behavior as a group. This could be accomplished, for instance, by significantly changing the costs of obtaining a particular set of characteristics or with the introduction of new bundles of the same types of characteristics.

Disaggregate choice models are classified as behavioral models in that they are consistent with aspects of both psychological theories

of choice behavior and the economic theory of consumer behavior. From the resulting model structure, a causal relationship between a particular choice environment and the consistent choices of individuals will be demonstrated.

The probability of choosing a particular alternative is related to an individual's consideration of his choice environment where that consideration is influenced by his own personal characteristics. Specification of a utility function to this end must take into account characteristics of the commodity in question and the characteristics of the individual,

$$U_j^a = (z_j, s^*[a]) \tag{2.1}$$

where U_j^a is the utility function of individual a over the jth alternative, z_j is a vector of attributes describing the alternative, and $s^*[a]$ is a vector of socioeconomic characteristics of individual a.

Different levels of utility exist for different ordered pairs of vectors of socioeconomic attributes and attribute vectors of alternatives.

2.2 THE PSYCHOLOGICAL APPROACH TO CHOICE: A BRIEF SUMMARY[6]

Psychological approaches to choice theory assume that choices are probabilistic and are based upon an evaluation of utility. It is assumed in particular that there is a direct correlation between the probability of a choice and the utility it gives to the individual. It follows from this that a ratio of the probabilities of two choices can be written as a ratio of the utilities of these two choices

$$\frac{P_1}{P_2} = \frac{U_1^a = (z_1, s^*[a])}{U_2^a = (z_2, s^*[a])}^{7}, \tag{2.2}$$

where P_1, P_2 are the probabilities of selecting alternatives 1 and 2 respectively.

The literature addressing this approach to the analysis of individual choice emphasizes that the choices are therefore based on relative utility assessment, and that this becomes a major behavioral assumption about the individual choice process.[8] Perhaps it is better stated that the above asserts that relative utility completely characterizes the ratios of the probabilities of taking alternatives, which would be, in a sense, observable.[9] Therefore, the implicit assumption of the approach is that utility is exact and measurable.

2.3 INTRODUCTION TO THE ECONOMIC APPROACH

The economic modeling of these choices differs from the psychological model on several of the aforementioned points.

1. The economic approach more distinctly emphasizes utility maximization as a major behavioral motivation of the individual. If the individual maximizes the level of wants satisfaction, and if the socioeconomic status of the individual is given, then this will be accomplished through the selection of alternatives which maximize $U(z_j, s^*[a])$, $j = 1, \ldots, J$.
2. More importantly, the economic approach through the emphasis on ordinal rather than cardinal utility leads one to question the adequacy of utility specification (2.1) in reflecting the nature of individual utility and choices.

We observe variation in choices over homogeneous groups of individuals. As the measurable and exact utility of the psychologists should reflect common choices for homogeneous populations, they must attribute this (often substantial) variation to imperfections in the perceptions of individuals of their choice situations. Economists prefer to modify the analysis to explicitly model unobserved differences between individuals in otherwise homogeneous groups, thereby giving theoretical interpretation to observed variations in aggregate demand.[10]

The utility specification is modified to

$$U_j^{a^i} = U(x_j, x_j', s[a^i], s'[a^i]) \tag{2.3}$$

where x_j represents a vector of observable and measurable attributes of alternative j, x_j' represents an unobservable vector of characteristics of alternatives (which are, however, perceived by the individual), $s[a^i]$ represents the vector of observable socioeconomic characteristics of individuals in group i, and $s'[a^i]$ represents a vector of immeasurable characteristics affecting the individual's tastes.

This more desirable specification reflects that the utility of the individual cannot be completely determined from observable data on socioeconomic characteristics of individuals and attributes of alternatives. Observed variations in preferences must be attributed to the unobserved characteristics which affect tastes. For that reason, the utility over any set of given observable characteristics will vary between individuals (of even the same socioeconomic group) in a stochastic manner. This model is formally known as the *Random Utility Model* as opposed to the *Strict Utility Model* asserted in the psychological

approach.

A generalized random utility U_j for the individuals in a population can be specified. Of course, any attempts to group individuals on the basis of their choices will be imperfect because of the effect of unobserved characteristics. Therefore, assume the Random Utility to be additive in two components.

$$U_j^{a^i} = U(x_j, s[a^i]) + \varepsilon(x_j', s'[a^i]) \qquad \begin{array}{c} U(x_j, s[1^i]) = U(x_j, s[2^i]) \\ \\ \varepsilon(x_j', s'[1^i]) \neq \varepsilon(x_j', s'[2^i]) \end{array} \tag{2.4}$$

where $U_j^{a^i}$ is the random utility of an individual in the ith group, $U(x_j, s[a^i])$ is observable and nonstochastic, and where U is measurable utility representative of the population. The value of the utility to the individual depends upon the values of x_j and $s[a^i]$. $\varepsilon(x_j', s'[a^i])$ is a stochastic portion of utility, where ε will be random even within homogeneous population subgroups.

It should be noted that although the utility is referred to as a Random Utility, all of the above arguments are known to the individual, and that there is nothing "random" about his own particular choice over alternatives.

The economic approach above leads to potentially superior utility specification than the Strict Utility Model, and also we are able to appreciate at this point Heckman's objections to the assumptions inherent in traditional continuous economic demand modeling. Traditional theory assumes a constant taste structure in the population, and that observed demand is distributed randomly about a common representative value. All systematic variations in demand must be interpreted as

common changes at the margin. This precludes recognition that the random elements of tastes cause distributions of choices throughout the population, and that these must be explicitly accounted for in any explanation of observed discrete choices.

If we postulate that an individual chooses between alternatives so as to maximize utility, then

$$P_1^a = \text{Prob}[U_1^a > U_j^a] \qquad j = 2,\ldots,J \tag{2.5}$$

[11]

where P_1^a is the probability that alternative "1" is selected by individual "a", and U_1^a, U_j^a are random utilities associated with the alternatives.

Substituting from (2.4):

$$P_1^a = \text{Prob}\left[U(x_1,s[a]) + \varepsilon(x_1',s'[a]) > U(x_j,s[a]) + \varepsilon(x_j',s'[a])\right] \tag{2.6}$$

$$j = 2,\ldots,J$$

which can be rewritten as

$$P_1 = \left[\varepsilon(x_j',s'[a]) - \varepsilon(x_1',s'[a]) < U(x_1,s[a]) - U(x_j,s[a])\right] \tag{2.7}$$

$$j = 2,\ldots,J.$$

The stochastic elements of utility favoring alternatives $j = 2,\ldots,J$ are compared to the observable (measurable) elements of utility favoring $j = 1$, which could be considered a constant $\alpha = \left[U(x_1,s[a])-U(x_j,s[a])\right]$.

A cumulative distribution function $F(\alpha)$ can be described for the stochastic portion of utility:

$$F(\alpha) = \text{Prob}\left[\varepsilon(x_j',s'[a]) - \varepsilon(x_1',s'[a]) < \alpha\right], \quad j = 2,\ldots,J. \tag{2.8}$$

From this a probability density function can be derived which, under certain assumptions, will lead to a particular structure of the probability of selecting a particular alternative j.

To this point, discussion has emphasized the individual. However, it should be noted that these probabilities will really represent frequency ratios of choices over particular alternatives for the population. For instance, if one supposes a homogeneous group i and then further records the choices of each member over perhaps two alternatives, then one will have obtained the frequencies of choices favoring alternatives 1 and 2. If the individual is chosen at random from this group, the selective probability described above is the probability that the individual will be a member of one of the designated groups. Therefore, as emphasized in equation (2.8), *it is the distribution of the unobservable characteristics which determines the probability that a certain individual will choose a given alternative.*

2.4 DERIVATION OF THE LOGIT MODEL

The derivation of a particular structure for the individual selection probabilities from (2.8) requires that a plausible distribution be postulated for the stochastic elements of utility. Several distributions can be assumed, each one yielding a different distribution of probabilities for the population, and therefore a different structure for the selection probabilities of the individual.[12]

One distribution often proposed for this purpose is the Weibull distribution.[13] The Weibull distribution is similar to the normal distribution except that it is skewed slightly to the right. As it is

not a particularly exotic distribution, the grounds justifying its use are basically in terms of computational convenience. It is attractive because of computation properties which seem to be consistent with some of the underlying processes implied in the model to this point.

One of these assumptions is that individuals make choices on the basis of the maximization of utility. The Weibull distribution is known to be stable under the process of maximization--that is to say that the value of utility under the process of maximization is known to be a random variable of a similar class as the variables over which the operation took place.

Another assumption is reflected in the form of the cumulative distribution function of (2.8). The distribution function has been expressed such that the random variable in question is the difference between the stochastic components of the random utilities of the alernatives facing the individual. The difference between two Weibull distribution random variables is a logistic distributed random variable.[14] Therefore, the distribution function (2.8) suggests that in the derivation of the choice probabilities for the individual, there will be a transformation from Weibull distributed random components of utility to logistic distributed probabilities over the same alternatives. This transformation is reflected in the structure of the final choice probability for the individual. The probability of a particular choice will relate to the cumulative distribution function of the logistic distribution $F(\alpha) = 1/1+e^{-\alpha}$.

The definition of the Weibull distribution is

$$\mathrm{Prob}[\varepsilon_i < \omega] = e^{-e^{-\omega}} \quad [15] \qquad (2.9)$$

Defining the Weibull distribution for the stochastic elements of utility yields the solution above as follows:

Let $\alpha_1 = [U(x_1,s[a]) - U(x_2,s[a])]$.

Let ω_1 = probabilistic value of error for $\varepsilon(x_1',s'[a])$

ω_j = probabilistic value of error for $\varepsilon(x_j',s'[a])$, $j = 2,\ldots,J$, where $\varepsilon(x_1',s'[a])$ and $\varepsilon(x_j',s'[a])$ are related to ω_1,ω_j through the relation $e^{-e^{-\omega}}$.

Simplifying the problem to a choice between only two alternatives:[16]

$$P_1 = \int_{\omega_1=-\infty}^{\omega_1=+\infty} \int_{\omega_2=-\infty}^{\max\omega_2=\alpha+\omega_1} \frac{\partial e^{-e^{-\omega_1}}}{\partial\omega_1} \frac{\partial e^{-e^{-\omega_2}}}{\partial\omega_2} d\omega_1\, d\omega_2$$

$$= \int_{\omega_1=-\infty}^{+\infty} \int_{\omega_2=-\infty}^{\alpha+\omega_1} e^{-\omega_1} e^{-e^{-\omega_1}} e^{-\omega_2} e^{-e^{-\omega_2}} d\omega_1\, d\omega_2$$

$$= \int_{\omega_1=-\infty}^{+\infty} e^{-\omega_1}\, e^{-e^{-\omega_1}}\, e^{-e(-\alpha+\omega_1)}\, d\omega_1$$

$$= \int_{\omega_1=-\infty}^{+\infty} e^{-\omega_1}\, e^{-(1+e^{-\alpha})}\, e^{-\omega_1}\, d\omega_1$$

$$= \int_{\omega_1=-\infty}^{+\infty} e^{-\omega_1}\, e^{-\log(1+e^{-\alpha})}\, e^{-e^{-\omega_1}}\, d\omega_1$$

$$= \frac{1}{1+e^{-\alpha}} \int_{\omega_1=-\infty}^{+\infty} e^{-\omega_1}\, e^{-e^{-\omega_1}}\, d\omega_1 \qquad (2.10)$$

$$P_1 = \frac{1}{1+e^{-\alpha}}\,(1) = \frac{1}{1+e^{-\alpha}}\,. \qquad (2.11)$$

The choice probability in the multiple choice situation follows from (2.8) by redefining the Weibull distribution for the multinomial case.

$$\text{Prob}[\varepsilon_j < \omega] = e^{-\sum_{j=1}^{J} e^{-\omega}} \tag{2.12}$$

17

$$P_j = \int_{+\infty}^{-\infty} e^{-\omega}\, e^{-\sum_{j=1}^{J} e^{-\omega_j+\alpha}}\, d\omega$$

$$= \int_{-\infty}^{+\infty} e^{-\omega}\, e^{-e^{-\omega_j}\left(\sum_{j=1}^{J} e^{\alpha}\right)}\, d\omega$$

$$= \int_{-\infty}^{+\infty} e^{-e^{-\omega}} \sum_{j=1}^{J} e^{-\alpha}\, d\omega$$

$$P_j = \frac{1}{\sum_{j=1}^{J} e^{-\alpha}} . \tag{2.13}$$

If $\alpha = U(x_1,s[a]) - U(x_j,s[a])$, then for a particular alternative in a multiple choice situation

$$P_1 = \frac{e^{U(x_1,s[a])}}{\sum_{j=1}^{J} e^{U(x_j,s[a])}} . \tag{2.14}$$

For simplicity and later reference, rewrite $\alpha = U(x_1,s[a]) - U(x_j,s[a])$ as $G(x)$ to reflect the combinations of observable variables influencing observed choices. Therefore, in the binary case,

$$P_1 = \frac{1}{1+e^{-G(x)}} \tag{2.15}$$

and in the multinomial case,

$$P_1 = \frac{e^{G(x)}}{\sum_{j=1}^{J} e^{G(x)}} . \tag{2.16}$$

2.5 IMPLICATIONS OF THE DERIVED MODEL STRUCTURE

2.5.1 The Cumulative Distribution Function

The relationships between the cumulative probability of a particular choice and the associated set of explanatory variables G(x) can be depicted as a sigmoid curve.[18]

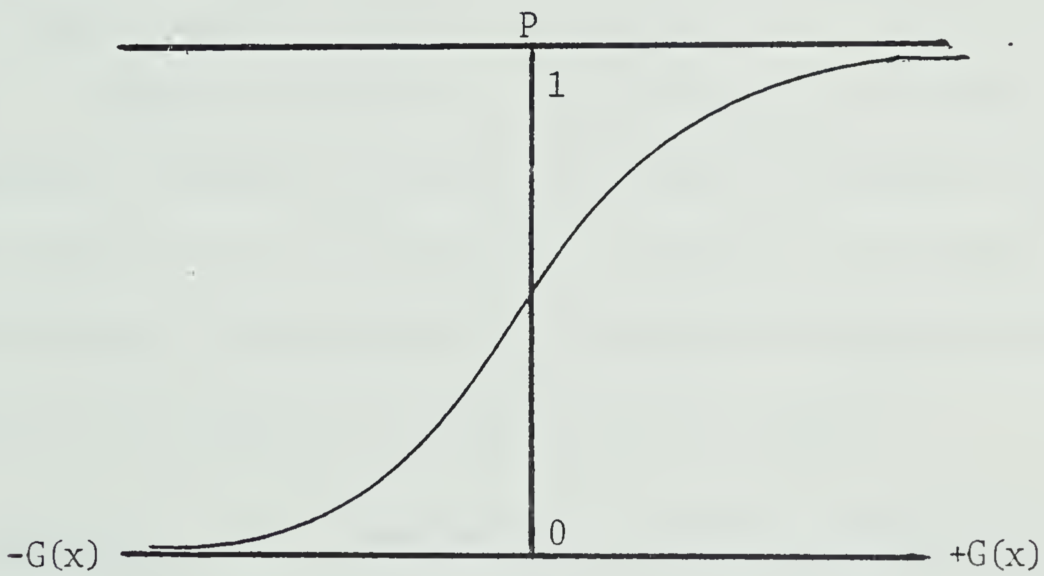


Fig. 2.1 The cumulative logistic function.

For any observed value of G(x) the associated probability is the frequency with which the individuals in a population will choose alternative j. In terms of the individual, it is the probability that he would have been in a group choosing alternative j.

Three observations can be made about this probability distribution:

1. The probability of choosing a particular alternative increases as the observed value of G(x) increases.
2. The probability increases faster as the values of G(x) increase through values around zero, i.e., if the alternatives are quite similar with respect to observed attributes so that differences in utility levels over observed attributes approach zero. If groups are observed with either high positive or low negative values of G(x), then it is possible to predict reasonably accurately the choice probability for that group even without precise measurement of the values of G(x). For individuals with values of G(x) close to zero, it is more difficult to predict accurately a particular probability of an action, as the cumulative probability changes more significantly for even small changes in the values of G(x).[19]
3. Fig. 2.1 emphasizes that it is the probability of a choice over a particular alternative which is the object of analysis, and that this probability is necessarily restricted to a 0-1 range. Therefore the logit transformation has served to transform even extreme values of explanatory variables to the 0-1 range.[20]

2.5.2 The Axiom of the Independence of Irrelevant Alternatives

The multinomial generalization of the model is

$$P_1 = \frac{e^{U(x_1,s[a])}}{\sum_{j=1}^{J} e^{U(x_j,s[a])}}, \quad j = 2,\ldots,J. \qquad (2.17)$$

Within this structure the odds of a choice between two alternatives within a multiple choice situation are derived through division.

$$\frac{P_1}{P_2} = \frac{e^{U(x_1,s[a])}}{e^{U(x_2,s[a])}} \; . \tag{2.18}$$

Although we derive this result through a model based on the premise of Random Utility, the result of (2.18) is identical in structure to (2.2), which was asserted in the Strict Utility Model.[21]

Both (2.18) and (2.2) imply that the existence of any third alternative has no effect on the ratios of the probabilities of choosing two given alternatives. The probabilistic ranking of alternatives implies that the Axiom of the Independence of Irrelevant Alternatives is satisfied. The Axiom states that if any two alternatives are available, and a third alternative is introduced, the ratios of the original probabilities will be preserved. In both (2.18) and (2.2) the ratios of the probabilities for $j = 1,2$ are independent of the existence of any third alternative.

One implication of the Axiom is that the introduction of new alternatives will reduce the probability of selecting *each* of the previous alternatives by the same relative amount, thus preserving the ratio of the original probabilities.[22] In the aggregate this implies that the cross elasticities of the market shares of the existing alternatives with respect to the introduction of new alternatives must be equal,[23] and that the share of the market going to the new alternative results in an equal percentage reduction in the market shares of each of the existing two alternatives.[24]

The implication of the Axiom of the Independence of Irrelevant Alternatives is subject to controversy. The psychological approach asserts that it is a major characteristic of the choice of the

individual. This might be seen in certain modeling situations to be an unreasonable assumption.

In the Random Utility Model, the structure of the model giving rise to the Axiom was derived through assumptions represented in (2.5) and (2.8), and the Weibull distribution for the stochastic element of utility. Some restrictions implying the validity of the model were implicit in the derivation. These restrictions are tantamount to requiring that only relevant alternatives to the individual be specified within the model and that the attributes of those alternatives be carefully and properly specified. This would, of course, be desirable for any modeling procedure, and therefore does not necessarily unduly jeopardize the range of the applicability of this procedure. At the very least, adherence to the restrictions implied in the derivation of the model ensuring the validity of the Axiom controls the potential misapplication of the model.

One of the restrictions was introduced in (2.4) with the requirement that $x_j \neq x_j'$--the stochastic component of the random utility cannot depend on the same attributes of alternatives as the nonstochastic portion of that utility. Without this restriction we might observe from the analysis that consumers have substantially different valuations of observed attributes, but that this would be due also to reasons related to unobserved tastes.[25] This could arise if the values of observed attributes were dependent upon other choices made by that individual. For instance, the individual makes choices which relate to the choice setting, thereby indirectly determining his own alternatives. The individual makes a choice unrelated to the direct choice between two alternatives, but that decision indirectly influences the observed

choices between the alternatives in question.[26] All of the choices would be incorrectly attributed to the value of the observed attributes of the alternatives. The result would be inaccurate calculations of elasticities between the probability of a particular choice and changes in the values of observed attributes.

Correlation between observed and unobserved attributes could also result from considerations related to the supply side of alternatives. For example, the existence of certain observed characteristics may induce competition from the producer of a competing alternative, but the terms of the competition may be to gain advantage in terms of unobserved attributes. Clearly, in this case, the existence of unobserved attributes is related to the existence of certain observed characteristics.

In both examples outlined above, the selection of a particular alternative cannot be separated from the existence of what might seem to be relevant substitutes--as they are not offered to consumers as independent alternatives--the result being that the Axiom will be violated and the model results would not be valid for that situation.

The Axiom will also be violated if there is correlation between unobserved attributes across alternatives. Therefore, $x_1' \neq x_j'$, $j = 2,\ldots,J$, for the same individual. The attributes must be clearly alternative specific variables. Complying with this restriction ensures that alternatives are relevant and correctly specified. This definition of a relevant alternative is not restricted in the Strict Utility Model. The result might be that two alternatives which are perceived by the individual to be near or perfect substitutes are treated in the model

as if they were distinct alternatives. In this case, the predictions of the model would be questionable.[27]

If, in a discrete choice problem, the above conditions are met, then the model would likely be suited to the problem. All systematic effects would have been identified as observed attributes, and the stochastic elements of utility would in fact be random. (Failure to correctly identify the observed attributes would result also in a violation of the Axiom.)

Skepticism arises regarding the validity and applicability of the Multinomial Logit Model due to the inherent assumption of the Axiom and the ease with which this Axiom can be violated. However, the conditions ensuring the validity of the Axiom are not unlike the conditions one would require for any well-specified model.[28]

2.5.3 Linearity of the Characteristics of Alternatives

The derivation of the logit model assumed the utility function to be linear in the exponent. The form of the choice probability in the binary model implies that consumers compare utility differences, and that these differences are linear functions of absolute differences in characteristics. McFadden [9] suggests that the specifications of the attributes can be complex transformations of the differences in attributes between alternatives.[29] The need for this would vary according to the type of choice being modeled.

It must be noted that linearity in the exponent function applies only to the attributes of alternatives, and does not apply to the socioeconomic characteristics of the individual. If socioeconomic characteristics are included additively in the utility function, then

they would disappear from the definition of the choice probabilities, as their values of the socioeconomic variables would be constant between alternatives.[30]

There are several ways of dealing practically with this problem.[31] However, it is not unreasonable to suggest at this point that socioeconomic characteristics do not necessarily affect choices additively, but instead interact in the valuation of the attributes of the alternatives being considered.

2.6 APPLICATION OF THE LOGIT MODEL TO THE TRANSPORTATION QUESTION

This thesis is specifically concerned with the application of the logit model to the analysis of transportation question and, in particular, questions related to the choice of mode for urban work trips.

2.6.1 The Need for Disaggregate Modeling in Transportation

Due to the urgency of urban transportation problems and the budget constraints facing urban governments, the following needs must be recognized:

1. Sensible and efficient policy to affect the behavior of individuals in existing transportation systems.
2. The ability to predict accurately the results of changes in transportation systems because of their effects on these choices.

Useful analysis requires accurate predictions of responses to particular policy variables in situations where choices at the individual level are discrete. This would suggest that these types of problems are amenable to disaggregate modeling. We have in fact observed an increase in the use of disaggregate modeling in

transportation in response to this need for prescriptive analysis, and many of the gains in this modeling technique have been achieved through their application to transportation problems.

The problems outlined in section 2.1 and Chapter 1 apply to the transportation case. To the planner, the importance of individual transportation decisions has been somewhat obscured by the more obvious aggregate traffic flows, and it has been this concentration by transport planners in the behavior in observed groups of travellers which has diverted attention from the level of the individual. As a result, traditional aggregate transportation models are incapable of examing the behavioral motivation behind transportation choices. These traditional models remain largely correlative with respect to observed traffic flows and characteristics of areas where these flows exist.[32] They cannot *explain* the relationship between those flows to potential changes in the characteristics of the system.

In order to achieve this a causal linkage between the characteristics of the system and the motivation for the choices must be modeled. This is the achievement of disaggregate modeling procedures over previous aggregate approaches. In addition, correlative studies of flows between areas are necessarily restricted to use in those areas actually used in the model, thereby reducing the general applicability of the results.[33] Disaggregate models, by concentrating on behavioral motivation, may yield results which are transferable between urban areas, or perhaps even through time.

2.6.2 Specific Transportation Demand Considerations

The demand for travel is a derived demand generated by the need

for accessibility to activities diffused over urban space. The decision to travel must be considered as a component of a larger overall communications system, where, given certain states of technology, even the decision not to travel is a viable alternative facing the individual. If the decision to travel is made, it is the simultaneous decision of route, timing, frequency and mode. Any one or a combination of these variables can be considered as substitutes for any of the other variables. Therefore, for an analysis of the overall travel decision, the choice set is not clearly defined. However, as is the case with any demand situation, the simultaneous estimation of choices from all available substitutes would be a complex problem, and also behaviorally would imply unrealistically broad perceptions on the part of the individual. The number of variables in the model cannot exceed the limits of human discrimination.

By modeling urban travel demand the problem is somewhat simplified, as in the short run many of the travel choices are demographically determined, thereby leaving fewer components of the choice to the discretion of the individual. This is especially true of the analysis of the relatively fixed travel patterns of the urban work trip. In the short run, the major discretionary choice left to the individual is that of mode choice.[34] The point of origin, destination, frequency and time of travel are predetermined by virtue of trip purpose, leaving the choice set more clearly defined. Also, the repetitive nature of the trip helps to ensure that the individual has well-developed perceptions of the values of relevant variables determining that individual's choices.[35]

The choice probabilities derived above can legitimately represent

the probability of taking a certain action within the overall decision to travel. These choice probabilities would be difficult to estimate if the attributes of relevant alternatives were to include variables representing all of the components of the transportation choice process. In the case of a binary choice, the comparative mechanism is in terms of the differences between the utility levels of different alternatives. The values of the attributes representing other components of the travel decision will be the same for the two alternatives and will then disappear from the specification. The choice probabilities will then be determined in terms of the differences between objective attributes of the alternatives themselves.

The multinomial model also lends itself to the estimation of choice probabilities for a subcomponent of the demand process. The structure of the multinomial choice probability does not restrict algebraically the attributes to those which differ between alternatives. All relevant alternatives, such as mode, time of day, route, etc., can be potentially included in the denominator of the choice probability. However, the Axiom of the Independence of Irrelevant Alternatives isolates the components of the travel decision. The probability of choosing a mode within a particular choice setting is independent of other options within the choice setting as alternatives.

One might consider the probability of a choice such as mode choice as a conditional probability,

$$P_j^a = \text{Prob}(m/t,f,d,\ldots) \tag{2.19}$$

where m = mode, t = time of day, f = frequency, and d = destination.

Satisfaction of the Axiom would imply that the conditional probability is really a marginal probability,

$$P_j^a = \text{Prob}(m) = \text{Prob}(m/t,f,d,\ldots) \qquad (2.20)$$

thereby justifying the modeling of a portion of the overall decision to travel.

Chapter 2

FOOTNOTES

[1]J. Heckman [16], p. 1.

[2]For a complete discussion of quantal response and extensive bibliography, refer to D. McFadden [31]. In particular, McFadden defined quantal choice as "an act that leads to various outcomes which can be indexed by finite or countable sets, where an individual is placed in one of N possible choice settings, and with each choice setting eliciting one of several responses from the subject." He further outlines that the response can be depicted as a drawing from a multinomial distribution of outcomes, each outcome having an "observable" choice probability. These probabilities are observable in the sense that it is possible to observe different frequency distributions of choices as the choice setting is repeated.

[3]That a choice setting is discrete is usually the result of the technology of consumption (marginal adjustments in quantity are infeasable) but can also be observed in consumption situations even when marginal adjustments can be made.

[4]Refer to D. McFadden [9] [31] for a survey and bibliography of this research.

[5]This is similar to the approach initially outlined by K. Lancaster [25].

[6]The basics of this approach are contained in Luce [26], Luce and Raiffa [27] and Stopher and Meyburg [50].

[7]The structure of choice probabilities for the individual for one mode can be derived from this. If $U(z_j,s^*[a])$ is specified as $e^{U(zj,s[a])}$, which is linear in the characteristics z, then a logarithmic relation in the operational model will result between the set of explanatory variables of alternatives and the selection probabilities.

$$\log P_j = U(z_j,s^*[a]).$$

One can rewrite (2.2) as

$$\frac{P_1}{P_2} \quad \frac{e^{U(z_1,s[a])}}{e^{U(z_2,s[a])}}$$

so that

$$P_1 = \frac{e^{U(z_1,s^*[a])}}{e^{U(z_1,s^*[a])} + e^{U(z_2,s^*[a])}}$$

and

$$P_2 = \frac{e^{U(z_2,s^*[a])}}{e^{U(z_1,s^*[a])} + e^{U(z_2,s^*[a])}} .$$

If the functions are linear with respect to the characteristics, then by dividing by $U(z_2,s^*[a])$

$$P_1 = \frac{e^{U(z_1-z_2,s^*[a])}}{1+e^{U(z_1-z_2,s^*[a])}}$$

$$P_2 = \frac{1}{1+e^{U(z_1-z_2,s^*[a])}} .$$

[8]As opposed, for instance, to a direct statement of utility maximization. In the case of cardinal utility these results are not inconsistent with those which would be derived from an assumption of utility maximization, i.e., $P_1 > P_2$ if $U(z_1,s^*[a]) > U(z_2,s^*[a])$. However, the maximization of utility does not *determine* the probability of a particular alternative being chosen.

[9]"Observable" as defined in footnote 2.

[10]Psychologists must attribute this instead to individual error.

[11]The superscript "i" will be suppressed for simplicity, but it should be remembered that these results hold for all individuals within observably homogeneous groups. Note also the difference in results from the psychological approach--the process of maximization is included in the probability of making a particular choice.

[12]Refer to T. Domenich and D. McFadden [9], pp. 56-62, for discussion of the specific implications of using different distributional assumptions for this purpose. They discuss the Probit Model (derived from the cumulative normal distribution), the Logit Model (from the Weibull distribution), and the Arctan Model (from the Cauchy distribution).

[13]This distribution has also been called the "extreme value" distribution, the "Gnechenko" distribution, and the "Inverse Exponential" distribution. See T. Domencich and D. McFadden [9], pp. 60-65, for detailed properties of this distribution.

[14]T. Domencich and D. McFadden [9], p. 63.

[15]M. Nerlove and D. Presse [36], p. 12.

[16]The basic steps for the proof are available in CRA [5], pp. 5-18.

[17]CRA [5], pp. 5-18, and the procedure is adapted from the proof found in Stopher and Meyburg [50], p. 281.

[18]The shape of such a distribution function is approximated by a variety of distributions, so this result is not unique to the use of a Weibull distribution.

[19]We would expect, then, that the effects of policy changes on the behavior of *population* will depend upon how individuals are clustered in such a distribution.

[20]See R. Pyndick and D. Rubenfeld [41], pp. 240-243, and T. Domencich and D. McFadden [9], pp. 58-59, for alternative methods of constraining the dependent variable to the 0-1 range if a Logit or Probit model is not used.

[21]Note, however, that the result in (2.18) includes only the observable characteristics of alternatives x_j and $s[a]$, whereas the result of (2.2) uses all characteristics of z_j, $s^*[a]$, assuming them all to be observable.

[22]For example, if a new third alternative has a probability of selection equal to .20, and the original probabilities for the existing alternatives were .50 each, then the selection probabilities for each of the existing alternatives become .40, thus preserving the ratio of the probabilities.

[23]If the selection probability for alternative 1 in a multiple choice situation is

$$P_1 = \frac{e^{G(x)}}{\sum_{j=1}^{J} e^{G(x)}} ,$$

then this probability with three alternatives

$$P_1 = \frac{e^{G(x)_1}}{e^{G(x)_1} + e^{G(x)_2} + e^{G(x)_3}} .$$

A cross elasticity between this probability and a change in a characteristic k of the third alternative is

$$E = \frac{\partial P_1}{\partial x_{3k}} \frac{x_{3k}}{P_1} = \frac{\partial P_1}{\partial G(x)_3} \frac{\partial G(x)_3}{\partial x_{3k}} \frac{x_{3k}}{P_1}$$

$$E = -P_3 \beta_k x_{3k}.$$

As reference to alternative 1 is not part of the result, this will also hold for the cross elasticity between the probability of selecting alternative 2, so that the cross elasticities are equal. Indeed, within a homogeneous population, we cannot, a priori, expect the cross elasticities to favor one alternative over the other. The information which would tell us this is unobservable.

[24]CRA [7], pp. D-7, emphasizes that this is only true for homogeneous populations.

[25]CRA [7], p. D-17.

[26]For example, if unobserved attributes of alternatives have an effect on individual location decisions. The new choice setting determines the values of the observed attributes (i.e., cost or time).

[27]In the case of transportation models, this is known throughout the literature as the Blue/Red Auto Problem. This often-used example concerns the existence of three modes of travel, two of which differ only in color, and are perceived by the individual to be identical commodities. Assume that $P_1 = 0.5$ and $P_2 = 0.5$ and that a third alternative is introduced which is a different color from alternative 1. Intuitively one would predict that $P_1 = 0.25$, $P_2 = 0.5$, and $P_3 = 0.25$, and that the consumer is indifferent between alternatives 1 and 3. The model will predict the probabilities to be $P_1 = 0.33$, $P_2 = 0.33$, and $P_3 = 0.33$ in order to preserve the original relationship of P_1/P_2. It is questionable whether in this case the model has led to acceptable predictions.

[28]CRA [7], pp. D128-D167 offer more extensive discussion of tests indicating violation of the Axioms and potential remedies for these problems.

[29]Domencich and McFadden [9], p. 54.

[30]It will be seen in Chapter 4 that this is a special case of the issue as to whether to use generic or mode specific variables in the model specification.

[31]Stopher and Meyburg [50], pp. 310-311.

[32]By concentrating on flows from one transportation zone to another, the analyst ignores differences between consumers within zones. These may be more varied than differences between zones. By aggregating information into measures of zonal averages, a large amount of information is lost.

[33]If no causality is established between the variables determining travel behavior and that behavior, then the models will be deterministic and useful only to those areas.

[34]See CRA [7], pp. C113-C161, for new developments in the modeling of other types of travel behavior.

[35]This is not to imply that individuals make certain choices out of habit.

Chapter 3

AGGREGATION AND TRANSFERABILITY

The application of new modeling procedures will ultimately be for policy purposes. The need to translate theory to policy recommendation raises two important concerns.

1. The advantage of disaggregate modeling has been in the gain in the ability to examine behavioral responses at the level of the individual. However, policy issues generally relate to the observed behavior of populations. It is necessary to aggregate the information regarding individual behavior in order to make predictions about populations.

   An explicit aggregation procedure must preserve the advantages of the disaggregate methodology over traditional aggregate models. If the model initially is able to demonstrate the sensitivity of choices to characteristics of the choice environment, then aggregate predictions involving the same individuals must also be sensitive to changes in these variables.[1] Naive aggregation procedures can mitigate some of the policy advantages of the disaggregate approach.[2]

2. If one goal of better modeling is to increase the analytical basis of policy, then it is necessary to recognize the real constraints of the policy maker in the use of quantitative analysis. Trade-offs often exist between new and better modeling techniques, and the resource requirements of their implementation. This has given

impetus to research on model transferability--the use of model results from one population in order to help make predictions about a different population.

A better understanding of transferability issues might mean that modeling procedures such as disaggregate analysis can be utilized at less cost to the policy maker, thereby increasing their adoption as regular procedures.

The discussion to follow in this chapter concerns the development of theoretically consistent aggregation procedures, and the delineation of preconditions for model transferability. It will be seen that it is the separation of the individual choice model (as described in Chapter 2) from the explicit aggregation procedure of this chapter which facilitates the potential for model transferability.

3.1 THE DEVELOPMENT OF MEASURES OF AGGREGATE BEHAVIOR

3.1.1 The Elasticity as a Summary Measure of Aggregate Response

One of the summary measures traditionally used to examine market behavior is the elasticity of demand. It is possible in this case to derive a continuous elasticity measure (even when the choices themselves are discontinuous) by using the probability of a particular choice as the dependent variable, and the explanatory variables determining the choices as independent variables. This elasticity measure can be stated as

$$E = \frac{\partial P_j^a}{\partial G(x)} \cdot \frac{\partial G(x)}{\partial x_{jk}} \cdot \frac{x_{jk}}{P_j^a} \tag{3.1}$$

$$= (1 - P_j^a) \cdot \beta_k (x_{jk}).^3 \tag{3.2}$$

Traditionally, elasticities represent market responses. In this case, the elasticity incorporates the choice probability of the individual, who is not in this case representative of the rest of the population. As the population actually represents a distribution of individual probabilities, an aggregate elasticity must take this distribution into account.[4]

3.1.2 Naive Aggregation Procedures

Naive aggregation procedures reflect traditional views of consumer demand. As similar consumers are assumed to have similar choices, a naive aggregation procedure uses the choice probability of a representative consumer to make predictions about the populations.

If this is to be an accurate representation of aggregate behavior, the group must be homogeneous. If this is true, then few aggregation problems would exist. One would merely calculate an expected probability for the population by using the mean values of the variables of the model.[5]

In a situation where there is a distribution of individual probabilities throughout the population (the cumulative response curve is nonlinear), the application of this aggregation procedure is clearly incorrect.

It is necessary to examine the entire distribution of probabilities in order to calculate an expected probability for the population.[6]

3.1.3 New Aggregation Procedures

Specific procedures used to take into account these concerns vary in terms of the accuracy of the resulting predictions. The actual selection of an aggregation procedure will involve trade-offs between

theoretical arguments and arguments based upon practicality.[7]

The most accurate procedure would be to examine the choice probabilities of each individual in the given population, and then to calculate the average probability

$$\bar{P} = \frac{1}{A(i)} \sum_{i=1}^{I} P_j^a \qquad (3.3)$$

where A(i) is the number of individuals in group i and P_j^a is the individual's probability of selecting alternative j.

This procedure requires a significant amount of information, which in some cases may not be available to the policy maker.

In some cases a more practical procedure may be to make use of summary measures of relevant distributions in the calculation of aggregate responses. Two options are available to the analyst at this point. One is to aggregate the model before estimating it, and then to use aggregated data to estimate the model. However, this results in a loss of information, and may yield conflicting results.[8] The other is to estimate the model in its disaggregate form, using individual data, and then to aggregate the results to whichever level is necessary to make the required policy predictions.

3.1.3.1 The aggregated model - a brief discussion

McFadden and Reid [33] have analyzed an aggregated model which makes use of the aggregate data, and attempt to indicate the conditions under which such a model is valid.[9] They use zonal averages for observed characteristics, as would a naive aggregation procedure, but they incorporate within the model necessary information about the

distribution of individual characteristics and attributes of the alternatives.[10] This accounting of the distributions of characteristics implicitly generates within the model structure an accounting of the distribution of individual probabilities.

Their results suggest that if populations are homogeneous, then the structure of the probabilities for the individual and the aggregate form of the probability are identical. The only difference would be that the aggregated form uses the mean values of reported variables to generate the probabilities rather than individual values. The advantage of this approach over traditional aggregate models is that the model treats homogeneous groups as a special case of many possible population profiles rather than imposing the assumption of homogeneous populations into the calculation of aggregate responses.

Although there is practical utility to this approach in that it is possible to use existing data more easily found in transportation data inventories, it fails to establish logical linkages between choice probabilities of individuals and observed aggregate responses.

3.1.3.2 Aggregate predictions from disaggregate estimates

With the advent of new modeling procedures, planning systems will change in order to accommodate them. The disaggregate approach implies a need for a new type of information inventory rather than just a different form of utilization of existing aggregate information. Although the aggregated model of McFadden and Reid is a short-run solution to the information problem, theoretical and statistical arguments favor the use of a disaggregate model structure estimated on individual data.[11]

Westin [60] clearly outlines the issues in modeling aggregate demand, and especially in the use of a logit model. He argues that aggregate results from even very simple policies are not obvious, and the results depend upon the distribution of the individual choice probabilities throughout the population.

The advantage of the aggregation procedure developed by Westin lies in its flexibility. His procedure is basically one of identifying a linkage between a known distribution of characteristics and a *derived* family of distributions for the selection probabilities of the population.

A certain number of minimum requirements exist in order that any derived family of distributions of probabilities will be useful in this type of procedure. Most importantly, the distribution of probabilities must be sensitive to changes in the underlying distribution of characteristics.[12]

Using the logit specification of individual probabilities of a particular choice, it is established from (2.15) that

$$P_j^a = \frac{1}{1+\exp^{-(\beta x_j)}} \ . \tag{3.4}$$

Westin describes the linkage from the characteristics of the population to a distribution of probabilities in the population in terms of transforming a distribution of population characteristics X to values of βX, and from this deriving a family of probability distributions.

Let $g(\beta X)$ represent a probability density function f(p) of (βX). If X is multivariate normal with vector means μ_k and a covariance matrix Σ, then βX will be univariate normal with parameters $\mu = \mu_k\beta$ and

$\sigma^2 = \beta\Sigma\beta'$.[13]

If

$$\ell n\left[\frac{P^a_j}{1-P^a_j}\right] = \beta X$$

then

$$f(p) = g\left\{\ell n\left[\frac{P^a_j}{1-P^a_j}\right]\right\} \cdot \frac{1}{P^a_j(1-P^a_j)} \tag{3.5}$$

and when $\left[\ell n \dfrac{P^a_j}{1-P^a_j}\right]$ is distributed univariate normal, this yields

$$f(p) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{P(1-P)} \exp\left\{- \frac{1}{2\sigma^2} \left[\ell n\left(\frac{P}{1-P}\right)-\mu\right]^2\right\}^{14} \tag{3.6}$$

This probability density function, previously described by Johnson [19] is of a family called the S_B family of distributions.[15] This family of distributions satisfies, as Westin notes, the pre-described conditions for a satisfactory derived distribution of probabilities. Importantly, as long as the normality of the characteristics distributions is preserved, the S_B family of distributions will always be derived, as it is this assumption which yields the structure of (3.6).

The link between the distribution of probabilities and the distribution of characteristics can be clarified through examination of the parameters of the S_B distribution, which, as seen above, are linear combinations of the distribution of characteristics.

One is again faced with two options in examining this relationship between the underlying characteristics and the distribution of

probabilities.

1. If the absolute magnitude of changes in the variables is known, the parameters of the transformed S_B function can be calculated (i.e., $\mu = \mu_k \beta_k$, $\sigma = \beta \Sigma \beta'$). The resulting changes in various moments of the derived probability distribution can then be directly compared. There is, however, some cost to this procedure, as the actual generation of the probability distribution and its various moments is a complicated procedure.[16]

2. It is less costly and a fairly straightforward procedure to examine the sensitivity of various moments of the distribution of probabilities to changes in the underlying characteristics. One of the moments most often analyzed is the expected probability E(P), which is defined as the expected proportion of individuals choosing a particular alternative.[17] Useful elasticities would then be in terms of the sensitivity of the expected proportion of individuals of taking an action to changes in moments of the underlying distributions of characteristics. Westin calculates two elasticities, generated from the S_B function, to be

$$E_\mu = \frac{\partial E(P)}{\partial \mu_k} \cdot \frac{\mu}{\mu_k} \cdot \frac{\mu_k}{E(P)} \tag{3.7}$$

where

$$E_\mu = \frac{\beta_k \mu_k E[P(1-P)]}{E(P)} \tag{3.8}$$

and

$$E_{\sigma^2} = \frac{\partial E(P)}{\partial \sigma^2} \cdot \frac{\sigma^2}{E(P)} \tag{3.9}$$

where

$$E_{\sigma^2} = \frac{E[P(1-P)]\{\ell n[\frac{P}{1-P}]-\mu\}^2}{E(P)} \; . \qquad (3.10)$$[18]

The information composing those elasticity measures is more readily available to the analyst, but is valid for only small changes in variable values.

This procedure represents a more sophisticated approach to aggregate predictions than the aggregate models used in traditional transportation planning procedures. The wealth of information which is available about individuals and their choices can be utilized to make reasonable predictions about group behavior. It can also be seen that these measures can be generated, if necessary, from only summary statistics describing populations. This will have important implications for transferability of results between models.

3.2 TRANSFERABILITY

3.2.1 Westin's Contribution Restated

Within the area of partial equilibrium modeling, it is often difficult to perceive how one can justify the use of model results to reflect demand considerations between diverse populations. That is to say, under the traditional theoretical partial equilibrium framework in demand analysis, it is difficult to account for differences between populations (which is important because the individuality of a group determines the range of observed choices) and yet, at the same time, reflect the common element of observed choice which might also be transferable. This is because the demand theory and aggregation usually coincide, and because the partial equilibrium nature of the

study precludes generality which might increase the flexibility of the model.

As implied in the introduction to this chapter, the methodology presented to this point while still partial equilibrium in nature, has achieved much toward a solution to this problem.

Westin's aggregation procedure utilizes the behavioral parameters of the model in conjunction with parameters of the distribution of characteristics within a population in order to derive a distribution of probabilities for that population. To this point in this thesis, and as has also been emphasized in the literature, Westin's contribution has been emphasized mainly in terms of its recognition as a more realistic and theoretically consistent aggregation procedure.

In addition, *the outline of a causal relationship between the characteristics of a population and the characteristics of a derived distribution of probabilities makes the aggregation procedure responsible for the sensitivity of predictions to differences in populations--freeing the model from this role.*

3.2.2 Research Results Addressing the Issue

Westin and Watson [61] indirectly provide some confirmation of this in their research into model transferability. They observe mixed success in their attempts to use model results between different origin/destination groupings of individuals, and thus attempt to find a systematic pattern and explanation for their success.

The aggregation procedure developed by Westin relies on both model results and information about the distribution of characteristics in the population. They are therefore provided with at least two

initial avenues with which to pursue their investigation.

1. The adequacy of the accounting of interpopulation differences within the aggregation procedure.
2. The coefficients of the explanatory variables, generated from the model.

The rationale for examining interpopulation differences as an explanation of their transferability success must be in the question of whether the aggregation procedure is sufficiently sensitive to interpopulation differences. The procedure utilizes summary measures of these distributions of characteristics. If these parameters are deficient in their representation of population differences, then they would expect to find poor results between the populations.[19] This would imply that it would be the aggregation procedure which would have to be upgraded to better take into account interpopulation differences.

After analysis of this question, they reject the notion that the observed diversity between populations explains their pattern of results. Therefore, for their purposes, the aggregation procedure seems adequate.[20]

Given the above, their pattern of success could possibly be explained in terms of the stability of coefficients of the models between populations. Only in cases where the coefficients are stable between populations can the coefficients of one population be used in place of those for a second population in order to make predictions about that second population. Upon analysis of this question, they find it necessary and sufficient for successful transferability that the model coefficients not be significantly different from one another.[21]

Superficially this seems to be a rather uninteresting result--

that one parameter can be used in place of another if they are equal. However, this result importantly suggests that the model has succeeded in identifying something about choice behavior which is common between populations.

3.2.3 Implications for the Tone of the Transferability Issue

If one interprets Watson and Westin's results in terms of placing equality of model coefficients across populations as a main determinant of transferability success, the main components of the transferability issue might be as follows: "*When, a priori, would one expect the model coefficients to be stable between populations?*" Related issues whose resolution would help clarify this would be:

1. What do the coefficients of the estimated model theoretically represent?
2. Under which conditions is it reasonable to theoretically assume their equality between populations? and/or
3. How are the coefficients theoretically affected by the assumptions of the model? (i.e., which assumptions, inherent in the structure of the traditional model, could be relaxed without affecting the stability of the coefficients?).
4. (a) How does the actual estimation procedure affect these coefficients? (i.e., within the estimation procedures, what can cause the coefficients to change?).
   (b) Are the calculated coefficients consistent with theoretical expectations?
5. Given the answers to these issues, within what limits should one expect equality of the coefficients?

An additional issue is whether one should adopt the Watson/ Westin result as the sole criterion for model transferability potential. Despite their attempts to explain their pattern of success in terms of the equality of coefficients, their actual criterion for transferability success is in terms of the quality of the predictions. Reasonable predictions may still be possible without equality of the coefficients.[22]

To concentrate solely on the above issues because of strict adherence to the Westin/Watson criterion might unduly restrict the use of the models in cases where the precondition of the equality of coefficients is violated, but where the quality of predictions might be reasonable.[23]

6. An important issue, then, is the sensitivity of aggregate results to changes in the values of coefficients, i.e., how different can coefficients be and still provide reasonable predictions between populations? If the aggregate results seem relatively insensitive to changes in the coefficient values, increased flexibility in modeling and estimation procedures could result.

Address to all of these issues is not attempted in this thesis. These could be viewed instead as important issues which must be addressed before a good understanding of the transferability issue can be achieved. The thesis will, however, directly address questions 1, 2, and 4(b), which will provide a reasonable basis for interpretation of the following empirical analysis and suggested answers to some of the other listed issues. In-depth pursuit of the other issues leads one outside the scope of the basic questions in this thesis.

3.2.4 Preconditions for Model Transferability

3.2.4.1 Interpretation of the model coefficients

From (2.4), the Random Utility takes the form $U_j = U(x_j, s[a]) + \varepsilon(x_j', s'[a])$. As $U(x_j, s[a])$ is the observable and measurable portion of the Random Utility, it is possible to identify a set of variables to characterize it.

Let

$$U_j = z_k(x_{jk}, s[a]) + \varepsilon(x_j', s'[a]), \quad k = 1,\ldots,K. \tag{3.10}$$

[24]

One is not able to specify variables for $\varepsilon(x_j', s'[a])$ as it is assumed to be known only to the individual. The stochastic portion of the utility affects the model, not through specification but through its effect on the structure of the choice probability.

Isolate

$$U(x_j, s[a]) = \beta_1 z_1(x_{j1}, s[a]) + \beta_2 z_2(x_{j2}, s[a]) \;.\;.\;.\; \beta_K z_K(x_{jK}, s[a]) \tag{3.11}$$

where z represents the particular transform of the variable to be used. This is actually the specification of G(x) in previous reference and which will be transformed to a probability through $P_j = 1/1+e^{-G(x)}$. The Logit is therefore described as

$$\ell n\left[\frac{P_j^a}{1-P_j^a}\right] = \beta_1 z_1(x_{j1}, s[a]) + \beta_2 z_2(x_{j2}, s[a]) \;.\;.\;.\; \beta_K z_K(x_{jK}, s[a]). \tag{3.12}$$

Initially, it is apparent that the coefficients (β_k) represent the change in the log of the odds of the choice of a particular alternative to a change in $z_k(x_{jk}, s[a])$, i.e.,

$$\frac{\partial\left[\ell n \dfrac{P_j^a}{1-P_j^a}\right]}{\partial z_k(x_{jk},s[a])} = \beta_k . \tag{3.13}$$

It is perhaps more meaningful to enquire of the sensitivity of the logit to changes in the value of a particular characteristic x_{jk}

$$\frac{\partial\left[\ell n \dfrac{P_j^a}{1-P_j^a}\right]}{\partial x_{jk}} = \beta_k \frac{\partial z_k}{\partial x_{jk}} \tag{3.14}$$

where $\partial z_k/\partial x_{jk}$ is dependent upon the form of the specification of the particular attribute in the utility specification. In this case, β_k represents the weight of that particular value in $U(x_{jk},s[a])$--the total set of variables describing the observable portion of utility. One observes an indirect link between this weight and the actual choice probability.

In the interest of further interpretation, one might enquire as to the relationship between the coefficients and the utility of the individual. The existence of a transferable model is dependent upon the assumption of a representative portion of the utility function for all individuals. Further insight into the implications of the stability of coefficients might be afforded.

If $U_j = (x_{jk},s[a]) + \varepsilon(x_j',s'[a])$ and $U_j = \beta_k z_k(x_{jk},s[a]) + \varepsilon(x_j',s'[a])$, then calculating the marginal utility:

$$\frac{\partial U_j}{\partial x_k} = \frac{\partial U}{\partial z_k(x_{jk},s[a])} \cdot \frac{\partial z_k(x_{jk},s[a])}{\partial x_k} + \frac{\partial \varepsilon(x_j',s'[a])}{\partial x_k} \tag{3.15}$$

$$= \frac{\beta_k(\partial z_k(x_{jk},s[a]))}{\partial x_k} + \frac{\partial \varepsilon(x_j',s'[a])}{\partial x_k} \,. \tag{3.16}$$

Two observations are in order:

1. If the stochastic and nonstochastic portions of utility are independent (which was previously stated as a precondition for non-violation of the Axiom of the Independence of Irrelevant Alternatives), then

$$\frac{\partial \varepsilon(x_j',s'[a])}{\partial x_{jk}} = 0 \quad \text{and} \quad \frac{\partial U_j}{\partial x_{jk}} = \beta_k \frac{\partial z_k(x_{jk},s[a])}{\partial x_{jk}}$$

is the marginal utility of a change in an explanatory variable.

2. If, in addition, the form of the specification of the variables is such that

$$\frac{\partial z_k(x_{jk},s[a])}{\partial x_{jk}} = 1,$$

then β_k represents the marginal utility of a change in an observable variable x_k. Only in this case is the value of the marginal utility independent of supply side considerations, or of existing specific levels of socioeconomic variables. Only under these conditions would marginal utilities of change in observed characteristics be equal between individuals.

3.2.4.2 <u>The stability of coefficients between populations</u>

The immediate interpretation of the above may seem at first to be

of questionable value.

1. Neither the forecasting procedure of Westin nor the results of Watson and Westin's research depends upon whether the coefficients represent marginal utilities or weights.[25]

2. Regardless of the independence between stochastic and nonstochastic portions of utility, *theoretically* the coefficients are related only to the nonstochastic portions of utility (i.e., should dependence exist, one could imagine some α_k to characterize

$$\frac{\partial\varepsilon(x_j',s'[a])}{\partial x_{jk}}).$$

Conceptually, under the assumption of a representative portion of a utility function for all individuals, the β's will always be equal between individuals and therefore populations for any given set of specified variables.

Importantly this is not to imply that one can expect stability of the coefficients between any two estimated models. The models are estimated on the basis of observed (reported) choices. In that case it is not clear that the actual coefficients will reflect the necessary restriction of representing only the nonstochastic portions of utility.

The issue, restated, becomes: "*Under which conditions can one expect the coefficients to reflect only the nonstochastic portions of utility, so that their stability between populations can be expected?*" One might view modeling restrictions for transferability as attempts to conform to this requirement. This also clarifies why analysts often state that successful transferability is the ultimate test of model specification.

One can view this question in terms of the following considerations:

1. Representation of the nonstochastic portion of utility presupposes that the results are independent of changes in supply side conditions.
2. From the previous discussion, it is apparent that no dependence between the stochastic and nonstochastic components of utility can exist.

A note about each of these considerations is warranted.

(a) Demand Side Isolation

The coefficients must reflect only demand side phenomena. If dependence exists between estimated coefficients and supply side considerations, then their stability across diverse populations cannot be assumed. The purpose of the procedure would be defeated if one was forced to arbitrarily assure similarity of supply side considerations in order to justify transferability of model results. Observable changes in supply side conditions will be reflected through explicit inclusion of the LOS variables in the specification of observable utility functions, and the results will be in the capitalization of these variables into the choice probabilities of the individual. The aggregation procedure will then combine this information about the individual and specific supply and demand side distributional considerations to generate predictions about the population.

The validity of this procedure (as is true for any partial equilibrium model) is dependent upon the assumption that the supply of characteristics is exogenous. This is not entirely unreasonable, especially in the provision of public transportation services--where

system characteristics are most often explicit decisions by public authorities. To the extent that the individual supplies some of the input to the transportation service (especially in the auto mode), this assumption could be violated. Many of the characteristics which are not determined by public authorities are determined by past household decisions. These, however, are fixed in the short run. Clearly these are short-run models.

Further restrictions exist so that the coefficients reflect demand side valuations related to the preferences of individuals.

1. Individuals should face unrestricted choices between the alternatives offered.
2. A reasonable set of relevant alternatives must exist.

Both of these considerations can be summarized under the idea of captive choice--where actual choice is eliminated because of the existence of some external constraint. This is usually corrected by explicitly accounting for these constraints in the model specification or else in the selection of individuals in the analysis.[26]

(b) The Axiom of the Independence of Irrelevant Alternatives

All coefficients in the model are estimated on the basis of observed 0-1 data, where the choices themselves are dependent upon both the observable and unobservable portions of utility. For good model transferability there can be no systematic relationships between these components.

From Chapter 2 it is apparent that dependence can exist because of:

1. The inclusion of irrelevant alternatives in the analysis.
2. Supply side competition which affects LOS variables (or is based

upon LOS variables) when the competition may be related to unobserved conditions.

3. A related but not previously discussed problem concerns the equilibrium of the household (or city). If the household is in a state of disequilibrium (i.e., with respect to other household decisions such as housing location), then changes in the LOS variables which are assumed to affect only transportation choice will actually have broader effects. For example, the utility may be affected, not just over that particular choice but, for instance, vis-à-vis other individuals. These effects are properly categorized as unobserved tastes, but which would be related to observed LOS variables, and which might affect their decisions. The modeling procedure therefore assumes that the area under study is in equilibrium, which could affect conditions under which one might expect transferability to hold. Considerations such as previous household mobility could be one such consideration.
4. The use of alternative specific variables in the specification may capture unidentified effects of taste for that particular alternative. Where one may find nontransferability of alternative specific variables, one may find transferability of generic variables.

Further observations can be made regarding transferability and violation of these preconditions.

Firstly, it was previously noted that if the Axiom is violated, the coefficients will be biased--for instance, when the coefficient captures all choice effects of that variable even when some of the effect is due to unobserved tastes. We would not expect these variables

to be stable across populations.

The biased coefficients resulting from such a situation will be used in forecasting procedures. Under certain considerations, the use in forecasting of these coefficients may be better than the use of corrected coefficients.[27] Such would be the case if one was able to assume the prolonged existence of a particular type of dependence throughout the forecast period. But this relates to particular idiosyncracies of a particular population, and it is unlikely that the same type of violation would exist between populations. In general, the transfer of these coefficients cannot be justified.

Secondly, no models, regardless of the suitability of the underlying conditions for potentially good transferability results, will ever be perfectly specified. This, along with problems associated with correcting for violations of the Axiom, might provide some rationale for updating procedures[28]--the use of new information to correct existing coefficients to reflect the new population for which they might be used. This, however, implies a need for at least small samples for the new population, which may or may not be available.

In conclusion, it must be noted that because of the case with which the preconditions for transferability might be violated, and the difficulties of correcting for this, research indicating the relative benefits and costs of these corrections is crucial. One might create more difficulties than one solves by searching for extravagant methods to conform to theoretical requirements for transferability. The coefficients are a means to an end, not the ends in themselves, so that the sensitivity of forecasts to small changes in coefficient values must be investigated. This information is important for determining the

increased flexibility in the modeling situation which can be allowed before compromising transferability.

Chapter 3

FOOTNOTES

[1]Changes in both the mean and variance of these characteristics distributions.

[2]Naive aggregation is a term coined by Westin and others to describe aggregation procedures which do not take into account the distribution of probabilities in a population.

[3]
$$E = \frac{\partial P_j^a}{\partial G(x)} \cdot \frac{\partial G(x)}{\partial x_{jk}} \cdot \frac{x_{jk}}{P_j^a}$$

$$= \frac{\partial \left| \frac{1}{1+e} - \beta x \right|}{\partial G(x)} \cdot \frac{\partial G(x)}{\partial x_{jk}} \cdot \frac{x_{jk}}{\frac{1}{1+e} - \beta x}$$

if $\frac{\partial G(x)}{\partial x_{jk}} = \beta_k$

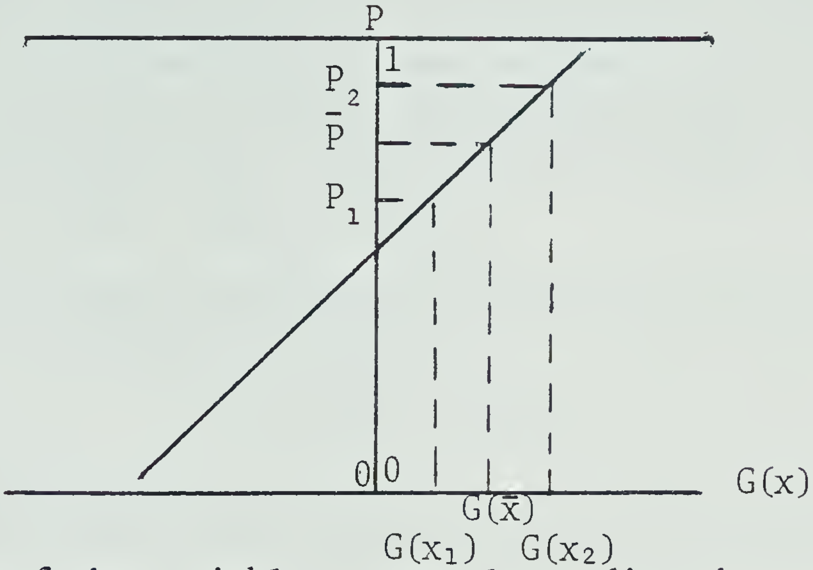
then $E = (P-P_j^a)\ \beta_k (x_{jk})$

$$= (1-P_j^a)\ \beta_k\ x_{jk}.$$

The elasticity is dependent upon the amount of the kth attribute possessed by alternative j, the weight of that variable in the utility function of the individual, and inversely related to the share of the market held by that alternative.

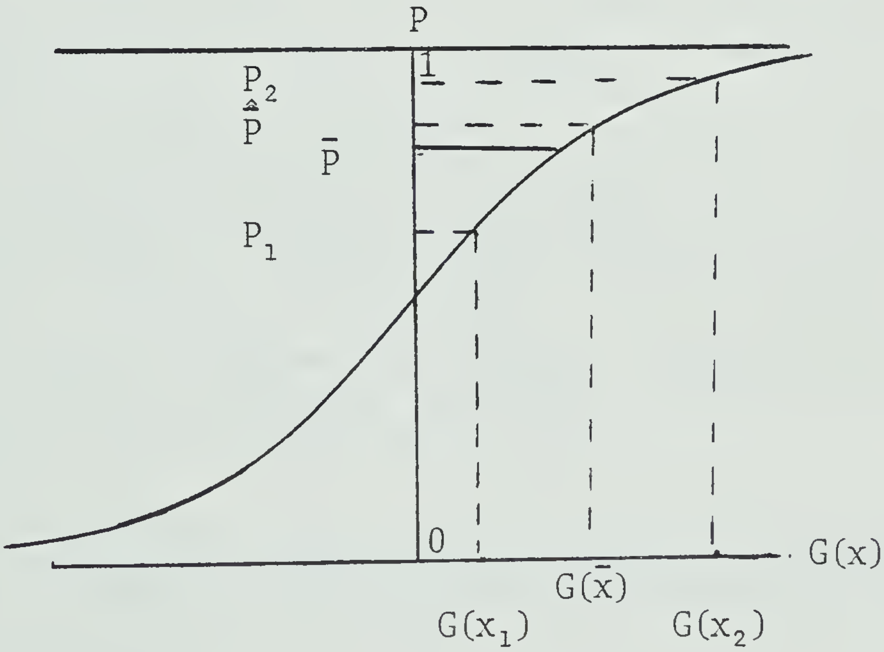
[4]P^a above is actually P^{ai}--an individual of type i. Domencich and McFadden [9], p. 50, outline and aggregate elasticity when market segments are identifiable. $E = w_i E_j$ where w_i is a weight representing the proportion of total demand for alternative j originating from people of type e, and E_j is their corresponding segment elasticity. They point out that the aggregate elasticity will be smaller than an elasticity evaluated at the population mean of the explanatory variables (naive procedure). This result is supported by Westin [60] who calculates an error $E(P)[1-E(P)]/E[P(1-P)]$ and by McFadden and Reid [33] who calculate that the aggregate elasticity is 82% of the individual elasticity.

[5]If there is no distribution of probabilities in the population, then the response curve is linear.



The mean values of the variables correctly predict the mean values of the probability ($\bar{P}$).

[6]In this example, the mean probability calculated from the mean values of explanatory variables will overstate the actual expected probability. In general, this estimated mean probability will be biased towards the closest extreme end of the distribution. This result is also important because the mean probability is used as an indication of the expected proportion of the population taking a particular action. Therefore E(P) itself is an aggregate prediction. See footnotes in Chapter 5 for these same estimates for this particular analysis.



[7]Koppelman [22], pp. 20-21, outlines a taxonomy of aggregation procedures classified in decreasing order in terms of data requirements and therefore also in terms of the accuracy of the reflection of the actual distribution of probabilities throughout the population.

[8]See Oum [39].

[9]For a full discussion of this approach, refer to McFadden and Reid [33]. This approach, despite its limitations, is an important contribution as it allows for the use of aggregate data which is already collected in traditional transport planning inventories, and which otherwise might be applied to less adequate traditional modeling procedures. 'Aggregate' data usually consists of zonal averages of transportation and user characteristics.

[10]Their individual probability using their notation and using a probit rather than logit specification is

$$P_j^{a^i} = \phi\ (\beta_1' z_k)$$

where ϕ is a standard normal distribution of $(\beta_1' z_k)$, z_k is a vector of differential attributes between modes, and β_k' is a coefficient vector. They find that

$$P_j^{a^i} = \frac{\phi(\beta' \bar{z}_k)}{\sqrt{1+\sigma_j^{i^2}}}$$

where $\bar{z}_k$ is a vector of mean values of differential attributes and $\sigma_j^{i^2}$ is the variance of the distribution of explanatory variables.

McFadden and Reid impose distributional assumptions upon the characteristics and therefore also the probabilities, using the assumption of normally distributed characteristics.

[11]Domencich and McFadden [9], p. 12, clearly establish that the intent of the modeling procedure is to outline differences in individual travelling behavior in an attempt to explain group behavior. It is theoretically and statistically desirable to examine a wide range of observations in making predictions about groups. If a large amount of variation in behavior is intrazonal, this should be preserved in the estimation of the model. Although McFadden and Reid's approach may reduce the costs of prediction in the short run, their approach suffers from some of the statistical disadvantages of other aggregate approaches. Long run costs could be reduced, for instance, by preserving the information on individual behavior variations, thereby reducing the size of required surveys and data inventories.

[12]See Westin [60], p. 3, for a specific discussion of these requirements. They lead to the underlying theme of his procedure--the identification of causation throughout the prediction procedure.

[13]Westin [60], p. 4.

[14]Ibid.

[15]Johnston [19].

[16]Westin [60], p. 9, notes that this actually required numerical procedures, which will likely be costly to the policy maker.

[17]The distribution of probabilities is importantly not equal to the distribution of outcomes, since the choice itself is a random variable. Inference, however, is made about population behavior from the distribution of probabilities. The 'expected probability' of individuals taking a particular action is inferred in terms of actual outcomes as the 'expected proportion' of the population of taking that alternative.

[18]Westin [60], p. 8.

[19]Note that the elasticity in (3.8) has the same implication as (3.2) but uses distribution information. Westin, R. and Watson, P [61], p. 236. For example, if the means or variances were not sufficiently sensitive to changes in distributions, thereby failing to transmit these changes through the aggregation procedure. Also, they measure 'success' through their attempts to predict actual observed outcomes, such as modal shares.

[20]Had they found that the model predicted well, for a given set of coefficients, between populations which were similar, but not between populations which were diverse, then it would imply that the aggregate procedure was not correcting prediction to account for these differences.

[21]They performed likelihood ratio tests in the null hypothesis that the coefficients were equal between population groups. They found that the hypothesis was rejected in cases where the predictions using these coefficients were poor, and accepted in cases where the predictions seemed to be adequate. In general, they found the coefficients to be different across models with different assumed trip purposes.

[22]This recognition is apparent in comments by both Westin [60] and Talvitie and Kirshner [21].

[23]This is important in view of the practical needs of policy makers who cannot afford (through time or data requirements) to further correct models to ensure transferability success according to the Watson/Westin criteria.

[24]McFadden [29].

[25]The coefficients are stated only, in most analyses, as representing a 'behavioral element'.

[26]These constraints include, on the auto side, the existence of legal means to use the auto, ownership of auto, and also includes familial constraints on its use. It has been empirically shown that 'auto availability' for the trip purpose being modeled is important. See CRA [7], Koppelman [22].

[27]Charles River Associates [7], p.

[28]See Atherton and Ben Akiva [2]. They refer to a process from Bayesion statistics. The estimated coefficients from a base model are

known, and in large samples have asymptotic normal distributions. The model is re-estimated on a small sample for the new area, which leads to a different distribution of coefficients. These are combined to get a new distribution of coefficients. They derive an updating formula for the single coefficient case.

$$\beta' = \frac{\dfrac{\beta_1}{\sigma_1^2} + \dfrac{\beta_2}{\sigma_2^2}}{\dfrac{1}{\sigma_1^2} \quad \dfrac{1}{\sigma_2^2}}$$

where β_1 = original coefficient

β_2 = new coefficient

σ_1^2, σ_2^2 = variance of coefficients β_1, β_2 respectively.

Chapter 4

SPECIFICATION OF THE UTILITY FUNCTION

The practical application of the Logit Model requires an appropriate representation of the utility function of the individual. Several characteristics of this utility function were established in Chapter 2.

1) The utility specification will contain variables of two classifications: those related to the socioeconomic characteristics of the individual and those related to the observed attributes of the alternatives. (These are commonly referred to as 'Level of Service' (LOS) variables.)

2) The logit $\ell n[P_j^a/(1-P_j^a)]$ is linear with respect to the characteristics of the individual and of the alternatives.[1]

3) Socioeconomic variables should not be included additively in the specification of a binary model except as mode specific dummy variables.

4. The Axiom of the Independence of Irrelevant Alternatives restricts the characteristics of the alternatives as being relevant and distinct between alternatives.

However, several other more general issues regarding the specification of this function must be addressed.

4.1 THE MEASUREMENT OF VARIABLES - GENERAL ISSUES

4.1.1 Alternative Specific Versus Generic Specification

This issue relates primarily to the method by which we describe a characteristic. Theoretically the concern is one of whether individuals have a common utility function for all relevant characteristics regardless of the commodities in which they are offered, or whether the valuation of characteristics does depend upon the commodity in which they are consumed (i.e., a different utility function exists for each alternative where the valuation of characteristics may differ between alternatives.)

An *alternative specific* variable[2] is one which is included in the specification of only one alternative, whereas a *generic* variable is one which is assumed to be valued equally in the utility function of alternatives. A generic variable can be considered to be a broad characteristic classification or "abstract" classification.[3]

Restrictions on the use of generic variables do exist. In the case of the calculation of probabilities in the binary logit model, variables which do not vary between alternatives will disappear from the specification--as the comparative mechanism is assumed to be in terms of the difference between modes. The Axiom of the Independence of Irrelevant Alternatives in the multinomial logit model will not be satisfied unless those variables included in a generic form have values which vary between alternatives. This limits the use of generic variables to Level of Service characteristics, as socioeconomic characteristics do not vary across alternatives.[4] Socioeconomic variables can be included generically if they are used in a form indicating interaction with a generic level of service variable. Behaviorally, this would imply that the variable influences a final choice only in its influence on the valuation of other variables.

Variables which vary across alternatives can be included as alternative specific variables. However, the use of alternative specific variables is restricted to situations where all individuals face the same set of alternatives.[5]

If socioeconomic variables are to be included in an alternative specific form, they are included as dummy variables, 0 for one alternative, and 1 for the other. These variables would then be interpreted as effects causing shifts in preference for the alternative in which it is specified.

Charles River Associates [7] have examined the issue of alternative specific versus generic variables in the interest of improving model specification for the purpose of increased model transferability, and for the projected use of the model for predictions regarding "new modes". The authors make the interesting point that the existence of different estimated parameters for variables between alternatives (indicating, then, possible validity for variables to be included in alternative specific form) might be interpreted as the effect of unidentified or unobserved characteristics which influence the valuation of those variables for particular alternative. They argue that "in a well specified model which explicitly accounts for all attributes that significantly affect choice, the use of a generic representative of Level of Service variables is justified."[6]

4.1.2 Additive Inclusion of Variables Versus Market Stratification

A choice must be made in terms of whether a variable should be included additively in the utility function, or whether the market should be segmented according to those characteristics. The issue is

one of whether variables are considered to have direct influence on final choices, or whether the existence of certain characteristics influence instead the valuation of all other characteristics, and therefore indirectly influence the final choice.

In this second case, each market subgroup would represent a different range of that independent characteristic. Behaviorally, this implies that behavior varies systematically across the strata. Different parameters for the variables across strata would indicate that individuals in different market segments value characteristics differently because of the attribute which they possess in common with other members of that strata. Therefore, one would expect that different market segments might be distinct in their travel behavior and/or their reactions to given changes in level of service variables.[7] This is useful information in a policy sense in that it gives information to policy makers regarding different reactions to policy changes from different market segments.

Market segmentation is common to marketing procedures.[8] To the extent that market segments are significant, identifiable and relate to either promotional strategies or service and planning options, transportation planners can be more effective in their policy.[9]

Stratification on the basis of socioeconomic or demographic characteristics is commonly found in the literature.[10] The variables are intended as proxies for culture or status which would affect the valuation of other variables.

Market segmentation might, however, be on the basis of variables other than socioeconomic variables, such as variables related instead to the supply side. Segmentation could also perhaps be on the basis of

attitudinal surveys.[11] Useful segmentation procedures might also be on the basis of variables reflecting choice constraints, i.e., auto availability, or alternatively on the basis of spatial considerations.[12]

Stratification on the basis of spatial considerations could possibly embody what normally might have to be a type of multivariate stratification, including implicitly some considerations such as supply side characteristics (and/or choice constraints) which tend to be homogeneous in regions of the urban place, some land use constraints, which influence both the supply side and underlying household decisions affecting travel behavior (i.e., home location), and possibly also could include relatively homogeneous groups with respect to demographic considerations.

One might try to approximate this by using concentric distance of trip origin from the trip destination as a basis for segmentation. Individuals would be homogeneous with respect to (straight line) distance for trips, and therefore have relatively similar cost patterns.[13] It is less clear that this type of stratification would represent homogeneous supply side or choice constraint situations unless the urban area is arranged concentrically in terms of land use.

4.2 ISSUES RELATED TO THE MEASUREMENT OF SPECIFIC VARIABLES

4.2.1 Level of Service Variables

Money costs and travel time associated with each mode are generally treated as two of the more important level of service variables affecting transportation choices.

(a) Money Costs

Some difficulties can be encountered in the description and

measurement of these costs, especially for the auto. One such issue is the choice between using perceived vehicle costs or an objective calculated per-mile cost.

Vehicle costs have traditionally included such out of pocket costs as gas, oil, parking and other tolls, plus a per-mile charge for maintenance and depreciation. One could easily conceive of a variety of other charges that the auto user encounters, including fixed charges such as capital costs and insurance. However, a distinction should be made between actual vehicle charges (either marginal costs or with the inclusion of some fixed charges) and vehicle and trip charges as perceived by the user.

Because of our emphasis on behavioral relationships, it seems reasonable that the model should use, where possible, those values for costs which the trip-taker perceives to be valid. These are the costs upon which the decisions are based, regardless of the correspondence of those costs to any number of calculated charges. Importantly, also, the model should use perceived *marginal* costs, or at least costs which can be most easily associated with individual trips.[14]

It has been shown[15] that auto users may underestimate costs other than the familiar out-of-pocket expenses. This gives added weight to the opinion that reported data on perceived costs should be used when available.[16]

The disaggregate modeling procedure requires data for all alternatives, whether or not they are actually selected. Due to survey inadequacies, reported data on the mode not selected is often not available, in which case calculated out-of-pocket expenses must be used.

The calculation of transit costs is straightforward, as actual

costs and perceived marginal costs are more likely to be the same.[17] The calculation of auto costs is usually a multiple of straight line distance from origin to destination.[18]

An important component of marginal trip costs is the parking toll associated with each trip. One faces the option of including either parking tolls with automobile charges and treating auto costs as total auto charges, or of treating parking as a separate commodity where the price is considered separately from other running charges.[19] The first option precludes recognition of parking choices as discretionary decisions separate from modal choices. The use of the second procedure would be in recognition that individuals may have separate valuations of charges associated with parking, and that it should be included as a separate level of service variable.[20]

One again faces the potential lack of reported parking information for non-auto users. However, average per-hour charges can be calculated either from the reported values of auto users or from the published fee schedules in the CBD from local parking authorities.

(b) Travel Time

An extremely important component of total perceived cost of travel of an individual by a particular mode is the opportunity cost of time involved in the use of each alternative. As individuals face different constraints with respect to time, the recognition of substitution in and out of alternatives on the basis of their time intensity is important.[21]

Some issues exist in the measurement of travel time for different alternatives.[22] Firstly, consider the use of measures of total access time for a mode versus measures of time involved on different segments

of trips on each alternative.[23] One can calculate a "value of time" by examining the apparent trade-off between time and cost in the utility function.[24] In using the value of total access time in the calculation of the value of time we would be asserting that the traveller has equal valuation of time saved across *all* segments of the trip. That is to say that in-vehicle time saved would be valued as highly as reduced walking time, and these would be equal to reduced waiting time. This is somewhat counterintuitive in that we might reasonably expect that an individual values time differently in different cases.[25]

Previous studies[26] support the procedure of segmenting total access time into the different segments of the trip. The most familiar segments are waiting time, in-vehicle time and walking time.[27] Segmentation is useful also in a policy sense if it can be substantiated empirically that individuals have different time elasticities in different trip segments, as policy makers presumably view time related attributes of different trips as distinct policy variables.

One is again faced with the alternative of using reported or perceived travel times of various segments versus engineering estimates. In practice, this will be determined by data availability. Ideally one would use reported estimates of time costs, as it is the individuals' perception of relative attributes of alternatives which determines their choices.[28]

c) Convenience and Comfort

Other level of service attributes can be classified into measures of convenience and comfort. Qualitative attributes such as these are difficult to identify and measure for the purposes of modeling.

The classification of travel time by trip segments can afford us

some interpretation in terms of the "convenience" of particular alternatives. Waiting, walking and transfer times are more likely to reflect supply side characteristics (especially in the case of public transit) than are such variables as in-vehicle time.[29] In-vehicle time would incorporate other influences such as the level of congestion, time of day or distance from origin to destination.

Waiting time, in particular, could be interpreted as a measure of convenience. We would expect waiting time to be least for individuals facing alternatives characterized by either high service frequencies or else consistent arrival times.[30] Alternatives with lower service frequencies and/or variable arrival times will likely report the highest waiting times, and therefore suffer the largest amount of inconvenience.

If information warrants, other measures of convenience and comfort can be used. Such variables might include the difference between estimated and actual time of arrival, or perhaps (for transit users) the availability of a seat for the duration of a trip.[31]

It is important in the calculated values for all level of service variables to attempt to approximate the conditions under which the user must travel. Importantly, for urban work trips, all calculated values for any level of service variables should reflect conditions at peak times rather than average daily conditions, so as to resemble perceived times as closely as possible.

4.2.2 Socioeconomic Characteristics

The choice environment for the individual includes both the level of service attributes of the alternatives and the influence of the

personal attributes of the individual, which determine his tastes and preferences. Potentially important personal characteristics include socioeconomic status, the stage in the family life cycle, and sex. From within these broad categories, several specific variables can be generated which might help to explain individual travel choices.

It has been argued that the proper variables to use in this context are those relating to individuals' total lifestyle, which is a combination of various kinds of personal attributes interacting together in the creation of certain behavioral motivations for choices.[32]

It can be argued that even socioeconomic measures can be a uni-dimensional view of lifestyle. The effects of economic affluence are important, but must be included with other variables indicating role differentiation, life cycle stage, and flexibility in travel decision making.[33]

In reading the literature and comparing applications of dis-aggregate procedures, one often comes upon a seeming contradiction. The theory determines that socioeconomic variables should not be included additively, yet almost all, including the best papers in the field, include at least some of the socioeconomic variables additively.

The theoretical derivation assumes homogeneous population groups. Actual model application will be over heterogeneous urban areas. Therefore, it is possible to parameterize these variables, and by doing so reduces the dependence of the LOS variables on these distributions of socioeconomic variables. Therefore, the restriction of nonadditivity of socioeconomic variables will be relaxed for practical application of this model.

(a) Income

Income as an additive mode specific explanatory variable is a purely economic orientation to socioeconomic status,[34] and as such could be viewed as an indication of budget constraint rather than necessarily a behavioral impetus to certain types of choices. (To the extent that higher income may induce more travel, it may have a more significant effect on nonwork mode choice.) However, when it is included in a model in an additive form, it is interpreted as a pure shift effect favoring a particular mode, showing a higher "preference" (all else equal) for that mode than individuals of a lower income.[35]

Use of the income variable in this fashion can be found throughout the literature, and CRA [7] establish empirically that the exclusion of income specific variables when in fact that behavior is important, will result in other parameter estimates which are dependent upon the income distribution found in the samples.[36]

One might wish to argue instead that income affects decisions through its influence on the perceptions of other decision variables (such as time or cost) rather than only having significant influence on its own.[37] This would lead one to stratify the sample on the basis of income, or define variables combining LOS and socioeconomic variables.

Income, in a practical sense, is often an awkward variable to use because of the difficulty in obtaining accurate data. Individuals seem much less willing to accurately report their income levels than other measures of their socioeconomic status.[38] Therefore if income is to be used as an additive explanatory variable, justification exists for the uses of ranges of income represented by codes, or even a simple distinction between high and low income individuals.

b) Life Cycle

An important individual factor affecting tastes and preferences is the stage in the individual (or family) life cycle. It is not unreasonable to assume that tastes and preferences change either as an individual changes age, or as family age, size and composition change. Variables reflecting life cycle can be included additively[39] or else logically could be used as a basis for stratification.[40]

Variables used as proxies for life cycle include age, marital status and number and age of children (although this latter variable could also reflect aspects of auto availability). The use of any of these variables is constrained by the availability of information, and their use in the literature ranges from the one of age as a dummy variable (indicating individual life cycle stage)[41] to family life cycle dummy variables reflecting composites of marriage, age and the number of children.

c) Role Differentiation

The choice of a particular alternative can be affected by the role of the user. This might be in terms of family priorities regarding the use of a particular alternative, or else because of the relative frequency of trips between one family member and another, necessitating the use of a particular mode.[42]

The variable most often used to reflect these considerations is sex, which is included in the model specification additively as a dummy variable. Considerations such as the above, however, can also be inferred from the interpretation of certain auto availability measures.

d) Auto Availability

The literature has shown that some measure of auto availability is an important determinant of travel behavior.[43] Justification for the inclusion of an auto availability variable could be in recognition of the difference between those captive on a particular mode (transit) and those who actually can choose between two alternatives. It could also be in recognition of the important equalizing effect that auto ownership has between two groups of otherwise different socioeconomic status.[44]

Measures used to reflect these choice constraints have been:

1) auto ownership (yes/no dummy variable)
2) number of autos/household
3) number of autos/licensed drivers
4) number of autos/worker
5) the use of both number of autos/licensed worker *and* number of autos/licensed nonworker
6) auto needed for work (yes/no dummy variable).

It has been argued [7] that measures such as auto ownership dummy variables or number of autos/household reflect only auto availability but do not address the problem of the availability of auto for the particular trip in question, and that these measures disguise the real problem of family competition for the auto at any one point in time.[45]

This is more closely addressed by the use of a variable such as number of autos/licensed driver or number of autos/worker, but these still neglect the question of whether competition exists for the use of that vehicle at a coincident time. In order to take into account both

competition from other licensed workers (which presumably reduces auto availability for any single user) and competition from licensed non-workers, it may be more expedient to use two variables, number of autos/licensed workers and number of autos/licensed nonworkers.[46]

The life cycle indicator of the number of children of certain ages (i.e., preschool) can also represent facets of auto availability. However, differing opinions exist as to the nature and direction of the effect. It may be the case that the existence of preschool children in a family reduces the mobility of that family, and therefore "frees" an auto for work trips (i.e., reduces competition for the auto). However, it may be true that preschool children are auto intensive (in terms of supposed impracticality of public transit use or reduced scheduling flexibility), thereby increasing nonworker competition for the auto.

The considerations listed above in the delineation of relevant variables for the model serve as the basis for the data selection and model specification for the analysis in Chapters 5 and 6 to follow. Actual model specification, however, will also be constrained by data availability.

Chapter 4

FOOTNOTES

[1]This implies a traditional common utility function with diminishing marginal utility for improvements in attributes. See Stopher and Meyburg [50].

[2]In transportation case this is known as a "mode specific" variable.

[3]Such variables would be time, cost, etc., in the general sense.

[4]Therefore the previous discussion of the restrictions on the use of socioeconomic variables is really a special case of the restrictions on the use of generic variables.

[5]CRA [7], pp. C14-C15.

[6]CRA [7], pp. C28-C31. They conclude this because the study does not disprove that time or cost is valued abstractly. This, they find, increases the potential applicability of these modeling procedures to the predictions regarding new modes.

[7]One such example is the documented insensitivity of high income groups to price changes. CRA [7], p. C9.

[8]See Transportation Research Record 649, "Preferences, Perceptions and Market Segments in Travel Behavior," for a variety of articles related to market segmentation. Particularly useful articles are Stopher, P. [47], and Nicholaidis, Wachs and Golob [37].

[9]Reichman [44] suggests five requirements for useful stratification procedures:
a) measurability
b) statistical robustness
c) substantiality
d) relation to travel behavior
e) relation to planning and service options.

[10]Typical variables used for stratification are income, age, auto ownership, and occupational group.

[11]For instance, stratification according to agreement scales, about congestion or attitudes towards LOS priorities, etc. Nicholaidis, Wachs and Golob [37] make a comparison of the use of demographic variables, choice constraints and attitudinal variables. They did find

that segmentations of the travelling population based on attitudes were found to have certain specific uses, but to be inferior to choice constraint segmentation for most planning purposes. Their work also includes a bibliography of other papers on the use of attitudinal surveys.

[12]See [37], [22] for study of the use of choice constraints for stratification.

[13]These costs include time costs, out-of-pocket costs, and the ratio of other charges (i.e., parking) to overall trip costs.

[14]Gasoline charges, parking tolls and other individual tolls can be considered to be marginal costs. One encounters, however, cost allocation problems with other fixed charges (i.e., insurance) between work trips and other discretionary travel. As these costs can be considered sunk, they have no effect on behavior on a trip-by-trip basis.

[15]Watson, P. [57].

[16]However, in recognition of the difficulty of measuring auto costs, CRA [7], pp. C81-C82, undertook a study to test the sensitivity of the model results to various specifications of auto costs. They found that "wide variations in assumed running costs do not significantly alter mode selection probabilities for disaggregate models."

Although the theoretical consistency of the specification of such variables is important, these results reduce, in a practical sense, the importance of the preoccupation with the correct specification of auto costs, and imply that effort should possibly be spent on the specification of variables when the specification is thought to have more influence on choice probabilities.

[17]The trip-by-trip transit fares will likely be the same, whether reported by the transit users or collected from transit fee schedules reported by transit authorities.

[18]An example of calculated costs used in [12] could be (on the basis of an average user at average speed)

$$C_i = \bar{c}\ (1.4)\ (D)$$

C_i is the cost of the ith trip
$\bar{c}$ is the average running cost of a car/mile
(1.4) = real distance/straight line distance
D = straight line distance.

[19]See D. W. Gillen [12] [13] for extensive discussion of this issue.

[20]As outlined by Gillen, the inclusion of parking charges correctly recognizes that there exists an effect on modal selection. However, the sensitivity of mode choice to changes in parking charges may be overstated if parking is included in a total cost variable. If, in fact, parking is a separate discretionary decision, reactions such as

parking relocation will reduce the observed effect on mode choice. Also, combined variables are less likely to be transferable to other areas.

[21]We observe individuals reacting to time constraints by arranging their consumption over time-related commodities according to the relative time intensity of each good. This is not inconsistent with traditional utility maximization where a consumer maximizes utility subject to a budget constraint, albeit in this case an expanded constraint including budget *and* time constraints. The consumer would then be seen to allocate consumption according to relative prices, but where the price now includes both time and money prices. In the context of discrete choice, all else equal, we would observe switching from one alternative to the other on the basis of significant differences in time intensities.

[22]And therefore potential measures of the value of time.

[23]Total access time might be considered as the sum of waiting time (transit), in-vehicle time, parking time (auto), and walking time.

[24]The Value of Time is the marginal rate of substitution between travel time and travel cost, keeping the consumer indifferent between two modes. CRA [5], pp. 48-49, provide an algebraic estimate of the value of time derived from a logit specification as β_t/β_c (INC), where β_t, β_c are coefficients associated with time and cost variables respectively, and (INC) represents the income of the individual.

[25]In particular, it is usually considered that waiting time is more onerous than either in-vehicle or walking time. This is substantiated by CRA [7], p. C89, "Travellers' response to travel time changes with the way time is spent."

[26]See Gillen [11], CRA [7], and McFadden [30].

[27]Walking time might alternately be included in a type of "total parking cost". Waiting time for public transit and "park and ride" modes might also include transfer time.

[28]Any engineering estimates of time for trip segments necessitate assumptions about the provision of the service. One technique for measuring waiting time for a public transport service is to use $\frac{1}{2}$ × (headway) for an alternative, regardless of the length of the interval. However, as indicated by Khalil [20], this is likely only to be an adequate approximation of waiting time for alternatives with high service frequencies. For longer intervals, individuals will likely arrange activities around an expected arrival time, in which case the procedure above will overstate waiting time. An alternative approach suggested is to use different formulas according to the headway of the alternative. A suggested formula from [1], [10] is 1.79 + 1.4 (headway) for intervals larger than 5 minutes, and $\frac{1}{2}$ (headway) for shorter ones. It has been shown [20] that improper estimates of waiting time have

large effects upon the calculated value of that waiting time, where these calculations range from 3 to 13 times the value of in-vehicle time, depending on which specification is used.

Calculation of travel time for in-vehicle, walking and transfer segments is basically straightforward. Walking time can be calculated on the basis of an average walking speed × distance walked. Transfer time would be based upon information given in published transit schedules and route guides of local authorities. In-vehicle time is usually calculated on the basis of a determined average speed multiplied by the straight line distance between origin and destination, or in a form similar to that of footnote #17. In-vehicle time for transit users would be similarly a sum of the average speed/link multiplied by the shortest route distance on each link.

[29]For instance, they could reflect such aspects of the supply as number of necessary transfer segments, frequency of service, reliability and the location of certain access points.

[30]Therefore individuals are able to arrange their activities consistently to minimize waiting time without the fear of "missing" the service.

[31]One might be able to use information on vehicle densities to proxy this measure of comfort.

[32]Reichman [44].

[33]Flexibility in travel decision making includes such aspects as control and timing (related to occupation groups or rough occupational classifications) and freedom in the choice of mode (auto availability).

[34]See Tardiff [52].

[35]See CRA [7], pp. C44-C49.

[36]CRA [7], p. C33.

[37]In fact, Tardiff [52] argues that if socioeconomic indicators are to be included additively, other socioeconomic measures such as education or occupation might be more adequate in that they offer more of a multidimentional view of lifestyle and its effects on behavior, and the effect of the income variable is reduced because of income related effects on other socioeconomic variables.

[38]This also might have implications for sampling procedures such as choosing between mail-in versus home interview surveys.

[39]Therefore being interpreted as a "shift" effect on preferences as was the case with the additive income variable.

[40]It is reasonable that certain LOS variables will be held in different priority for individuals and families of different ages.

[41]Age can also be interpreted as having an effect on social engagement patterns which in turn affect travel choices. One would expect this to be more significant in nonwork trip mode choice rather than work trip choice.

[42]The literature generally indicates that females take less trips than males; therefore sex = male being positively related to the use of auto mode.

[43]Nicholaidis [37], Koppelman [22], CRA [7].

[44]Reichman [44], p. 41.

[45]"Auto availability" as a variable can also be defended on the grounds that auto availability acts as an inducement to taking trips, but this also is likely more important for discretionary travel rather than work trip mode choice. It would only "induce" travel for the work trip to the extent that occupation allows *other* travel during working hours.

[46]One might expect that the importance of number of autos/licensed worker will be more important than number of autos/licensed nonworker as a measure of competition for vehicle. This is because of the obvious usual coincidence of working times. However, to neglect the number of autos/licensed nonworker would neglect the effect of competition for other auto use, which has still been found to be significant. See CRA [7], pp. C58-C59.

Chapter 5

DATA DESCRIPTION AND EMPIRICAL RESULTS

5.1 DATA DESCRIPTION

5.1.1 Sample Selection

Ideally, one would use individual reported data for estimation of the disaggregate model. Information on individual trip behavior, characteristics of trip makers, and reported LOS characteristics are available for the Ottawa/Hull region for the spring of 1975.

The data to be used for this analysis is a subset of a sample collected by the Transportation Development Agency in 1975, which was initially collected as part of an overall survey commissioned by the Ministry of Transport at that time.[1] The original sample was collected through return mail questionnaires, requesting information within the broad categories listed above. All of the individuals surveyed were employees of the federal government.

The specific hypotheses underlying the development of the questionnaire are documented, as is information regarding survey response rates, editing, coding and factoring procedures.[2] These will not be discussed in detail at this point. However, the hypotheses underlying the survey design are consistent with Chapters 2 and 4 of this thesis, and the data received from the original sample had been edited for internal consistency and completeness.

From this large sample, observations were selected which had

complete information on a number of important variables.[3] This sample of 934 observations included users of a variety of modes for work trips, including vehicle drivers, vehicle passengers, taxi, bus, motorcycle, bicycle and walking. A smaller sample of 827 auto drivers and bus users for the summer of 1975 was derived from this sample.[4]

As the data received from the Transportation Development Agency's original sample was subject to previous editing, only a few further adjustments were necessary.

1. Captive bus users within the city area were deleted. This was accomplished on the basis of responses to questions regarding auto availability for work and the ownership of a driver's license.
2. All respondents living greater than sixteen miles from work were deleted on the assumption that they lived out of the city. There was a large breaking point in the sample at this distance, with the next highest reported distance being 23.5 miles.[5] This is a generous distance allowance to fulfill an assumption of city location given the size of the Ottawa/Hull area, but as the direction of travel is not known, cross city travel must be taken into account.
3. Obvious remaining coding mistakes and clearly inconsistent cost reports were deleted (i.e., reported costs which were inconsistent with time and distance information).[6]

These three adjustments resulted in a loss of 60 observations, leaving a final sample of 767 users facing a choice between the use of the auto or public transit for work trips.

4. The analysis to follow necessitates stratification of the sample on the basis of distance from work zone to home. The sample was stratified by two procedures, with the number of stratifications in

each being arbitrarily equal to four. The first procedure (with subsamples denoted as models 1, 2, 3 and 4) reflects a stratification on the basis of a logarithmic scale. The logarithm was used as an arbitrary weight to equalize on the margin the burden of additional miles.[7] The second procedure (with subsamples to be denoted as models 1^*, 2^*, 3^* and 4^*) divided the sample into equal four-mile intervals. The first procedure resulted in more consistent sample sizes in each model than did the second procedure. The distance ranges (miles) in each subsample are shown in Table 5.1.[8]

TABLE 5.1

SUBSAMPLE DISTANCE RANGES

MODEL	MINIMUM	MAXIMUM	MODEL	MINIMUM	MAXIMUM
1	0.0	4.3	1^*	0.0	4
2	4.4	6.8	2^*	4.1	8
3	6.9	10.8	3^*	8.1	12
4	10.9	16.0	4^*	12.1	16
CITY	0.0	16.0			

5.1.2 Sample Description

(a) Area

All of the respondents in the final sample live in the Ottawa/Hull region. The area of Ottawa/Hull provides a reasonable metropolitan area for this study. The two major modes are clearly defined, i.e., there are no competing public transit modes in the area, so that public

transit reflects only one type of alternative. Also, because of the common employer of the respondents, the work areas are clearly defined and were edited in the original Transportation Development Agency sample to be in the CBD.

The use of the Ottawa/Hull region could potentially provide a problem for transferability studies if these were to be done between it and other metropolitan areas. One might wonder if by nature of being the national capital and the fact that a majority of employees are government employees, that the city is sufficiently atypical to affect the transferability of Ottawa results to other urban areas. This problem is eliminated in this particular case because the transferability will be tested within the Ottawa/Hull boundaries. However, as part of the transferability issue is in the observation of common preferences, one might observe better results in such a homogeneous city than in other urban areas.

One complicating factor in applying the following transferability tests to the Ottawa/Hull area is in the aggregation of these two centres in the sample. One would expect large English/French cultural differences to exist between the two, and cultural differences may, a priori, be expected to reduce transferability success.

As the spatial stratification in these models is in terms of concentric distance from the CBD, there will be a mixture of English/French respondents in each area model. Information confirming the cultural backgrounds of the respondents is not part of the sample, so it is impossible to determine the proportion of English and French in each subsample.

(b) Timing of Trips

All respondents were asked to report travel behavior for the work trip. Therefore, within the limits of flexible working hours allowed by the federal government, the time and cost values reported reflect common peak conditions.

(c) Mode

As stated previously, the different modes available for response on the questionnaire were clearly delineated. This was especially useful in the distinction between the two travel modes of "auto driver" and "auto passenger". All further reference to the "auto" mode refers to the auto driver rather than to the auto passenger.

The potential problem of defining the "bus" mode for respondents living outside of the city lends further weight to the decision to delete them from the sample. They would either be likely facing a different "bus" mode than would city dwellers, or would be captive auto users.[9]

(d) Auto and Bus Travel Time

All of the time values in the sample are reported perceived values for one-way travel. (It was recognized in Chapter 4 that the use of perceived LOS information is more consistent with behavioral modeling than is calculated data.) All respondents in the sample have reported time estimates for the mode not selected, which eliminated the need to calculate time and cost values for any individuals.

Reported auto time ranged from 2 minutes to 80 minutes, and bus time reports ranged from 3 minutes to 150 minutes. The mean reported auto and bus times for each distance range are indicated in Table 5.2.

TABLE 5.2

ZONAL MEANS FOR ONE-WAY AUTO AND BUS TIME

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Mean reported auto time	21.048	16.617	20.069	23.203	26.663	16.301	20.533	24.263	29.705
Mean reported bus time	39.016	30.890	39.261	41.988	47.692	30.688	37.912	45.263	53.689

TABLE 5.3

ZONAL MEANS FOR ONE-WAY AUTO AND BUS COSTS

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Mean reported auto costs (5¢)	15.776	12.947	15.015	16.454	21.308	12.430	15.422	17.689	22.787
Mean reported bus costs (5¢)	7.249	5.493	7.320	6.968	11.317	5.554	6.705	7.814	14.016

Previous analysis by De Leuw Cather [8] of the original sample from which this sample was drawn, compared perceived time values to calculated time values for auto and bus. They indicate that, in general, bus users underestimated bus time, auto users overstated auto time, and bus users highly overstated auto time.[10]

(e) Auto and Bus Costs

All cost information indicated in the sample is also perceived costs for one-way travel. Bus costs were predictably similar within a range of 25¢ to 40¢ for one-way travel. The mode response was clearly 25¢. A marked increase in reported bus cost is observed as one examines distances implying out-of-city location, and bus costs reported for distances to home zone greater than 16 miles from work approach $40.00 for one-way travel. Mean reported auto and bus costs for each distance model are indicated in Table 5.3.

Previous analysis comparing perceived and calculated cost values for the original Transportation Development Agency sample indicated that auto drivers understated auto costs and bus users much overstated them.[11]

(f) Walk Time

This information is partially available for auto users, and was specifically stated in the survey as "walk time from parking to office." This potentially valuable information was, however, incomplete even for the auto users, and a parallel question was not directed to bus users. A variable reflecting walk time will therefore not be used in the model specification.

(g) Parking Costs

Auto users are surveyed as to parking costs, but the question of potential parking charges is not addressed to users of alternate modes. The parking information is quite complete for auto users. These reported parking costs were classified according to the reported work zones of each respondent, and examination of these indicated that average parking charges did vary somewhat with the work zones. Therefore, parking cost values were generated for the remaining auto users, and all bus users. The charge ascribed to each was the average charge paid in their respective work zones. It is recognized that this could provide a misleading picture of parking opportunity in that it does not reflect parking availability. The averaging procedure used also reduces the range of parking charges used in the models relative to the actual variability in the sample. This makes it more difficult to discern the effect of parking costs on mode choice. Parking charges reported range from $0 to $55, with the average in the city-wide sample being $22.37.

(h) Sex

This information is available for the entire sample. The profile of the samples is given in Table 5.4. A higher percentage of females live near the city centre than live in the outer distance ranges. This probably reflects differences in family formation in each area, with career-oriented families or singles living closer to the CBD.

(i) Age

Age categories were reported in all cases, where the age categories on the survey were 25 and under, 26-35, 36-45, 46-55, and 56 and older. These categories were converted to two categories, one with

TABLE 5.4

NUMBER OF MALES AND FEMALES IN EACH DISTANCE RANGE

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Sex M	696	181	177	234	102	161	319	159	60
F (%)	71 (9.3)	28 (13.4)	26 (12.8)	17 (6.8)	2 (1.9)	25 (13.4)	34 (9.6)	9 (5.4)	1 (1.6)

TABLE 5.5

AGE PROFILE OF EACH DISTANCE RANGE

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Age x<26 x>55 (%)	98(12.8)	52(24.9)	50(24.6)	21(8.4)	20(19.2)	42(22.6)	70(19.8)	13(7.8)	7(11.5)
26<x<55	669	157	156	230	84	144	283	154	54

TABLE 5.6

INCOME PROFILE OF EACH DISTANCE RANGE

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Income < 5000	3	1	1	1	0	1	1	1	0
5000-9999	54	24	17	11	2	23	27	4	0
10000-14999	139	42	38	43	16	32	69	27	11
15000-19999	177	49	35	61	32	41	72	46	18
20000-24999	195	44	53	74	24	41	95	45	14
25000-29999	125	27	41	37	20	27	61	26	11
> 30000	72	20	18	24	10	19	28	18	7

age less than 25 years or else older than 56, and the other for age reported between 26 and 55 years (the age group for most workers).

The population profile in these two groups is shown in Table 5.5. There is a higher percentage of non "middle-aged" workers in the area models closer to the workplace than in the suburban zones of the study area.

(j) Family Size

Respondents were asked to indicate the number of children in the family in various age groups. Unfortunately, the information is very incomplete. This will leave the model specification deficient with respect to this proxy for both life cycle and auto availability.

(k) Income

Income group responses are available for all individuals in the sample. The responses are coded under seven categories: less than \$5,000; \$5,000-\$9,999; \$10,000-\$14,999; \$15,000-\$19,999; \$20,000-\$24,999; \$25,000-\$29,999; and greater than \$30,000. The income profile of the population is indicated in Table 5.6.

5.2 MODEL SPECIFICATION

From 3.5, the equation to be estimated is

$$\ln \frac{P_j^a}{1-P_j^a} = G(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_6 x_6 \qquad (5.1)$$

where $j = 1 =$ auto (A)

$j = 0 =$ bus (B)

and $x_1 = MIN = T_A - T_B$ (differences in travel time)

x_2 = COST = $C_A - C_B$ (differences in travel cost)

x_3 = PCOST = reported or calculated parking charges

x_4 = SEX (1 = male, 0 = female)

x_5 = AGE (1 = $25 < x < 55$; 0 = otherwise)

x_6 = INC (coded income groups from 1 to 7).

A ratio specification is also tested where MIN = T_A/T_B, COST = C_A/C_B and PCOST = PARK/C_B. Comparison of the results between the difference specification and the ratio specification indicates that the ratios are less appropriate for this study.

One would expect, a priori, that $\beta_1 < 0$, $\beta_2 < 0$, $\beta_3 < 0$, $\beta_5 > 0$, $\beta_6 > 0$. Given the lack of information in the model about family size and auto availability, it is more difficult to predict, a priori, the sign of β_4. One would expect $\beta_4 < 0$ if family constraints reduce auto availability for the primary worker and if this is a widespread effect. On the other hand, one might argue that if the male is a primary worker, and the auto is available, then $\beta_4 > 0$.

5.3 ESTIMATIONS AND RESULTS

5.3.1 Estimation Procedure

It has been established in Chapter 2 that logit analysis represents a procedure which estimates the probabilities for individuals of choosing a particular alternative. An observed sample of choices represents selected random variables resulting from logistically-distributed probabilities, reflecting Weibull-distributed random utilities.

The estimation procedure of Maximum Likelihood is conceptually consistent with this procedure. This estimation method determines a

set of parameters which maximize a function describing the likelihood of having observed a particular sample of random variables from a set of known distributions.

Maximum Likelihood estimation, which is well documented in many sources, yields estimates which are unique and have asymptotically desirable properties.[12] The model is estimated in this case using the Maximum Likelihood Procedure incorporated in the LOGIT ANALYSIS section of the SHAZAM collection of programmes.[13]

Estimation results will include the estimated parameters of the model, results indicating their statistical significance, the goodness of fit between predicted and actual outcomes, and measures of significance for the entire equation. Since the parameter estimates have asymptotic normal distributions, asymptotic t statistics are provided in the results. These test the null hypothesis that each of the parameters has a mean equal to 0.

A traditional measure of goodness of fit is the multiple correlation coefficient R^2. An analogous measure can be used for logit analysis,[14] but since the model parameters have only desirable properties in large samples such a measure is also reliable only for large samples. McFadden [29] [9] indicates that a more satisfactory measure can be obtained through the use of the log likelihood functions.[15] Greig [62] calculates a pseudo-R^2 measure equal to

$$\frac{1 - e^{(2 \log \lambda/N)}}{1 - e^{(2 \log Lo/N)}} .$$

McFadden [29] indicates that pseudo-R^2 measures of 0.2 to 0.4 indicate a good fit between predicted and observed action.

The likelihood ratio test indicates the significance of the entire equation. When the results of a likelihood ratio test are presented with estimation results, it is a test of the null hypothesis that all of the coefficients are equal to zero. The likelihood is calculated under the constraints indicated in the null hypothesis, and then without these constraints. The ratio of these likelihoods is denoted as λ, and it can be shown [53] that $-2\log\lambda$ is distributed asymptotically chi square with K^* degrees of freedom, where K^* is the number of constraints indicated in the null hypothesis. This test will also be used to test the equality of the coefficients between models--where the null hypothesis will be that the parameters are equal between different equations.

5.3.2 Estimation Results

5.3.2.1 The use of a difference rather than ratio specifications

Both the difference specification and the ratio form of the variables for time and cost have limitations. The difference formulation ignores proportionality in cost or time perceptions. A five-minute time saving is given equal weight between a ten-minute travel time and a sixty-minute total travel time. The ratio formulation, while allowing for proportionality in the perceptions of time and cost savings, could impose rather extreme assumptions about that proportionality. For example, a five-minute time saving in a total time of ten minutes is assumed to have twice as much weight as the same time saving for a twenty-minute trip. This assumption could also be restrictive.

The estimation results for the two cases are given in Table 5.7. These results are based on estimations on a city-wide model.

TABLE 5.7

COMPARISON OF MODEL RESULTS FOR A RATIO VERSUS DIFFERENCE SPECIFICATION FOR MIN AND COST

VARIABLE	PARAMETER	(t VALUE)	VARIABLE	PARAMETER	(t VALUE)
MIN = $T_A - T_B$	-0.08414	(-9.6100)[a]	MIN = T_A / T_B	-0.80120	(-2.8278)[a]
COST = $C_A - C_B$	-0.013299	(2.513)[b]	COST = C_A / C_B	-0.096702	(-2.8021)[a]
PCOST	-0.040275	(-1.7025)[d]	PCOST = PARK/C_B	-0.42353	(-5.8885)[a]
SEX	-0.33067	(-1.0644)	SEX	0.71452	(2.4089)[b]
AGE	-0.03239	(-0.14585)	AGE	0.14707	(1.9020)[d]
INCOME	0.19861	(3.2298)	INCOME	0.14409	(2.2368)[c]
CONSTANT	-0.45361	(-0.69665)	CONSTANT	1.0329	(1.7900)[d]
Pseudo-R^2	0.26502		Pseudo-R^2	0.14640	
Likelihood ratio test*	166.945	(6 D.F.)	Likelihood ratio test*	89.1906	(6 D.F.)

*Critical value of $-2\ell n\lambda$ at 99.5% = 18.5476

(a) significant at 99.5% confidence
(b) significant at 99% confidence
(c) significant at 95% confidence
(d) significant at 90% confidence

The results of the likelihood ratio test indicate that both models are significant and both models provide estimates of correct sign and reasonable magnitude. A decision to use the difference form of the model specification for the following analysis is based upon the relative significance of the two models, and because of the lack of a priori reason to select the ratio formulation despite the lower model performance indicated above.

5.3.2.2 Estimation results of individual models

The parameter estimates and test statistics for all models estimated are provided in Table 5.8. All of the models are significant at greater than 97.5% confidence. The signs of each LOS variable are in the expected direction in all models, as is true also of the income variable. There is considerable variation in the directions of the effects of the other socioeconomic variables and the constants, but these variables are not significant in any models. Time difference seems overwhelmingly to be the most important independent variable, determining choices in all models.

It seems also that the parameter estimates of the LOS variables are similar across all models, although this must be confirmed by more precise methods.

Once the parameter estimates for the model are known, individual probabilities can be generated through a transformation of the index $G(x)$ to the individual probabilities P_j^a. This transformation is achieved through

$$P_j^a = \frac{1}{1+e^{-G(x)}} .$$

TABLE 5.8

ESTIMATED MODEL COEFFICIENTS FOR NINE DISTANCE STRATIFIED MODELS

MODEL	City	1	2	3	4	1*	2*	3*	4*
DISTANCE RANGE	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	1-4	4.1-8	8.1-12	12.1-16
MIN	-.08414 (-9.610)	-.066828 (-3.833)	-.064944 (-4.3617)	-.12174 (-6.6214)	-0.082463 (-3.2028)	-.061867 (-3.4933)	-.083461 (-6.4399)	-.12686 (-5.5210)	-.064614 (-2.2261)
COST	-.013299 (-2.251)	-.012743 (-1.0192)	-.010331 (-1.5893)	-.018829 (-.95677)	-.017059 (-1.1576)	-.0041082 (-.30929)	-.021036 (-2.1668)	-.003482 (-.30795)	-.041872 (-1.7089)
PCOST	-.040275 (-1.702)	-.010979 (-.29912)	-.043080 (-.697231)	-.084810 (-1.8503)	-.056638 (-.59860)	-.003209 (-.090305)	-.051770 (-1.5075)	-.14965 (-1.9321)	-.17169 (-1.0316)
SEX	-.33067 (-1.0644)	.23860 (.52122)	-.82566 (-1.3159)	-1.3096 (-1.5119)	-.050066 (-.049163)	-.087914 (-.18748)	-.97824 (-1.9066)	-.096161 (.085138)	-.98289 (.72006)
AGE	-.03293 (-.14585)	.023619 (.066448)	-.058146 (-.14843)	.22605 (.41255)	.29398 (.37486)	.18823 (-.50436)	.049949 (.15230)	-.096932 (-.12321)	1.3283 (1.2472)
INCOME	.19861 (3.2248)	.23121 (2.2759)	.087825 (.70662)	.28771 (2.3300)	.20109 (1.0201)	.22977 (2.1759)	.21712 (2.2550)	.20422 (1.3340)	.12336 (.44634)
CONSTANT	-.45361 (-.69665)	-1.5755 (-1.5482)	.93794 (.75280)	.29250 (.20336)	-.51845 (-.19686)	-1.4444 (-1.4002)	.34234 (.35280)	.84963 (.36990)	3.0272 (.73926)
PSEUDO-R^2	.26502	.17524	.20460	.42657	.25781	.15249	.27730	.43337	.32004
LIKELIHOOD RATIO TEST 6 (D.F.)	166.945	29.2974	32.7233	95.2408	21.5108	22.4563	81.0231	63.4182	15.7439
# OBSERV.	767	209	203	251	104	186	353	167	61

Critical value for $-2\ln\lambda$ at 99.5% = 18.5476; at 97.5% = 14.4494

TABLE 5.9

SUMMARY STATISTICS OF THE ACTUAL PROBABILITY DISTRIBUTION FOR NINE DISTANCE STRATIFIED MODELS

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0.4	4.1-8	8.1-12	12.1-16
Binomial estimate	.6067	.5598	.6453	.5976	.6538	.5645	.6006	.6407	.6885
E(P)	.6066	.5522	.6358	.5976	.6539	.5645	.6006	.6407	.6885
E(P[1-P])	.19168	.21091	.19234	.16171	.18475	.21797	.19030	.15356	.1648
E[G(x)]	.61749	.26358	.75539	.78859	.93890	.28720	.58116	1.0649	1.1645

A logistic curve for each model can be generated which is the cumulative probability in each sample of taking the auto. These distributions are shown for each distance model in Figs. 5.1 through 5.9.

Summary statistics describing each distribution are presented in Table 5.9. These statistics include a binomial estimate for each model, which in this case is the *actual* number of auto drivers as a proportion of the total sample. This is analogous to the traditional estimate in a binomial distribution, indicating the number of "successes" as a proportion of a total number of trials of a particular test.

The *estimated* equivalent to the binomial estimate is the expected probability E(P) of taking the auto. This is the mean of the estimated distribution of individual probabilities generated from the model. Table 5.9 also includes the expected variance of the estimated probability distribution E[P(1-P)] and the mean values of the calculated indexes E[G(x)] for each model.[16]

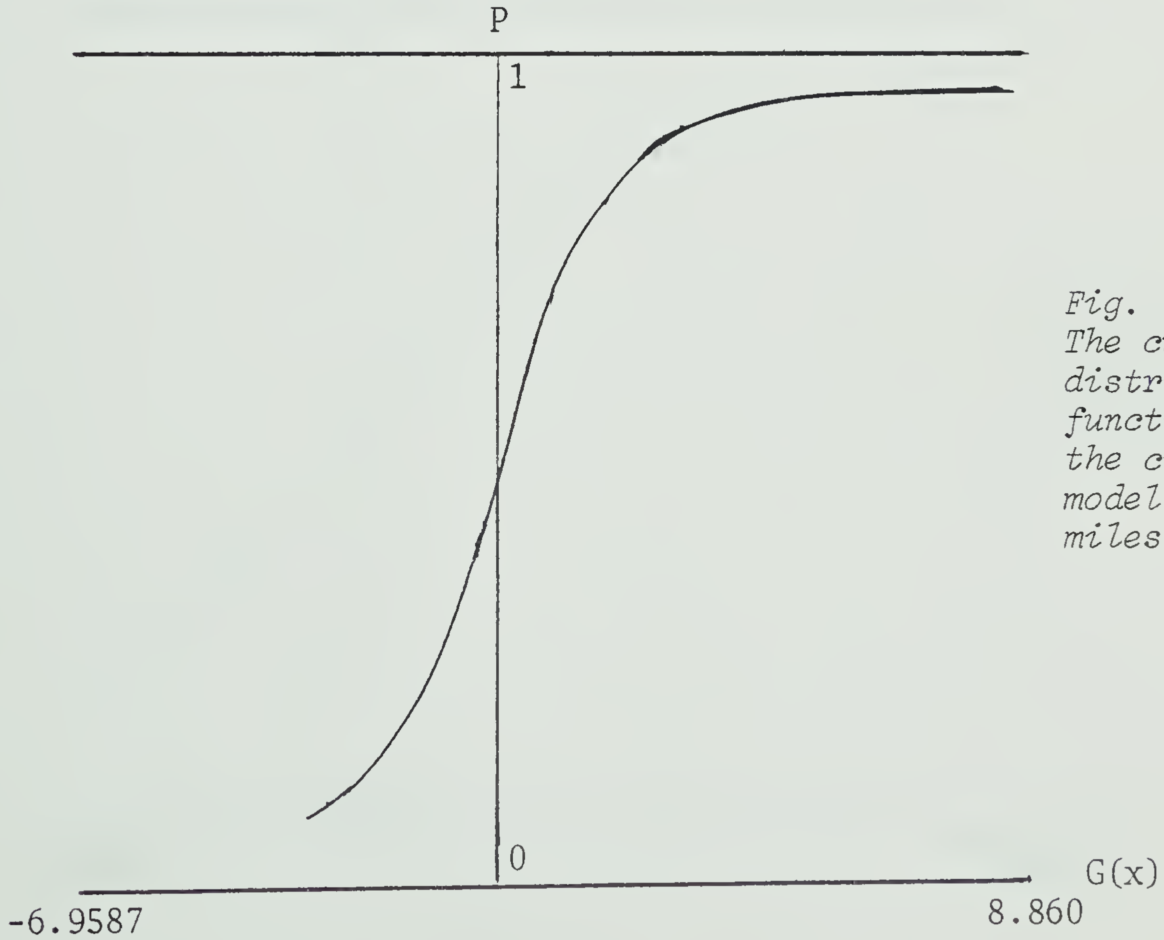


Fig. 5.1 The cumulative distribution function for the city-wide model (0-16 miles).

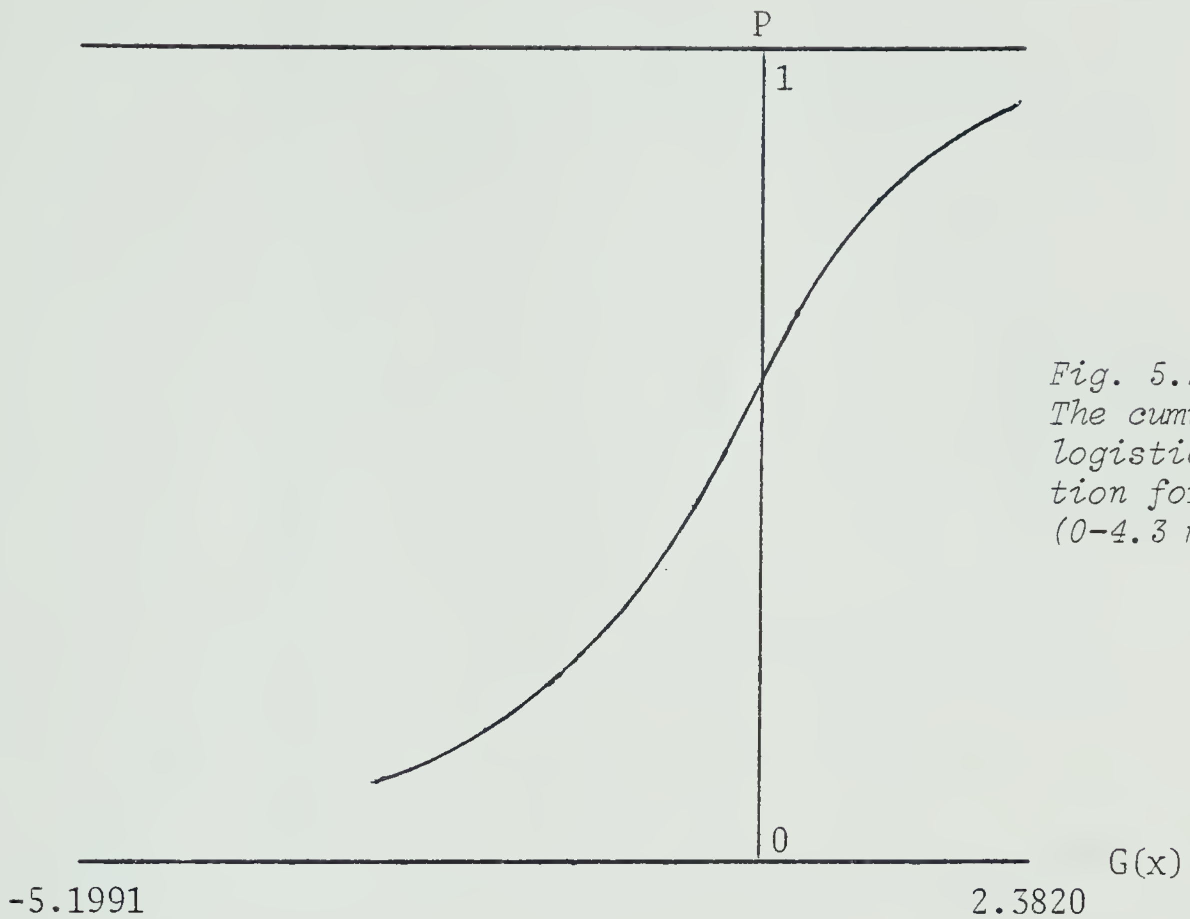


Fig. 5.2
The cumulative logistic function for model 1 (0-4.3 mi)

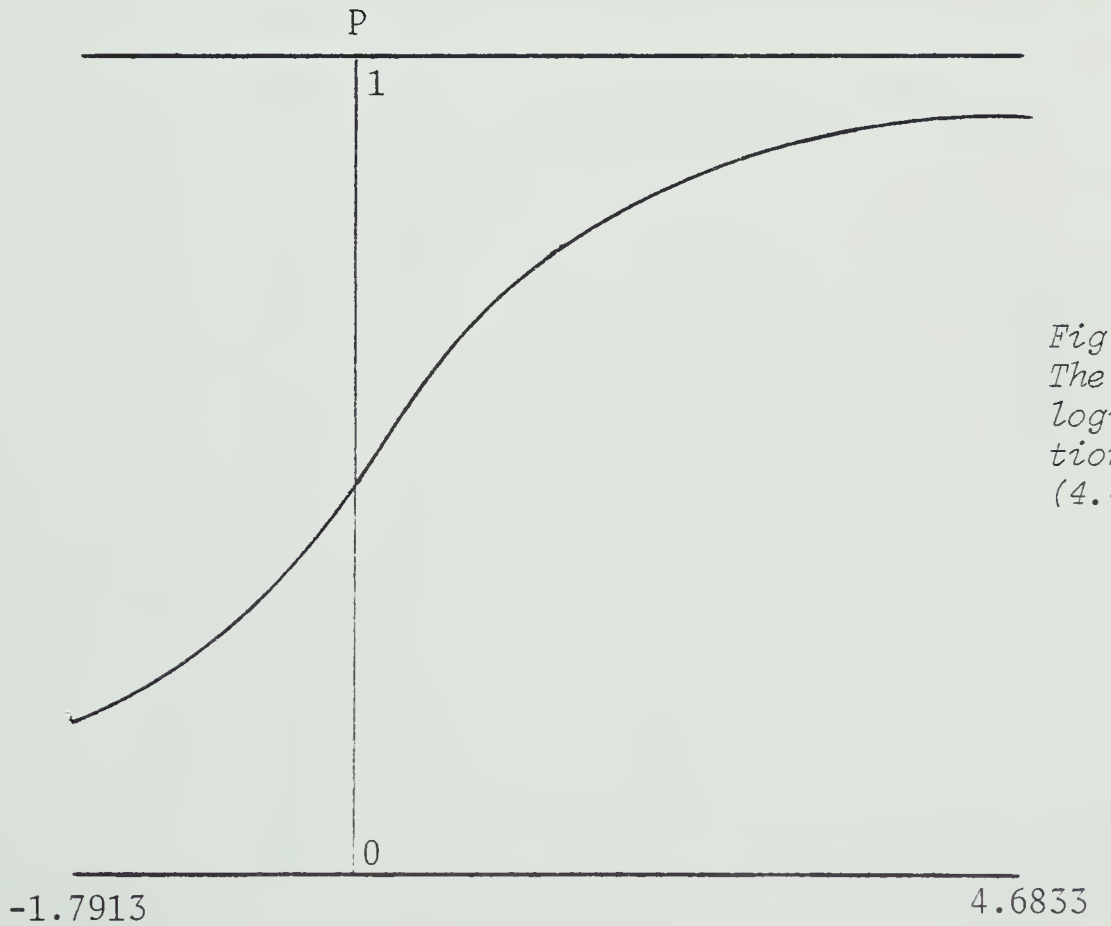


Fig. 5.3
The cumulative logistic function for model 2 (4.4-6.8 mi)

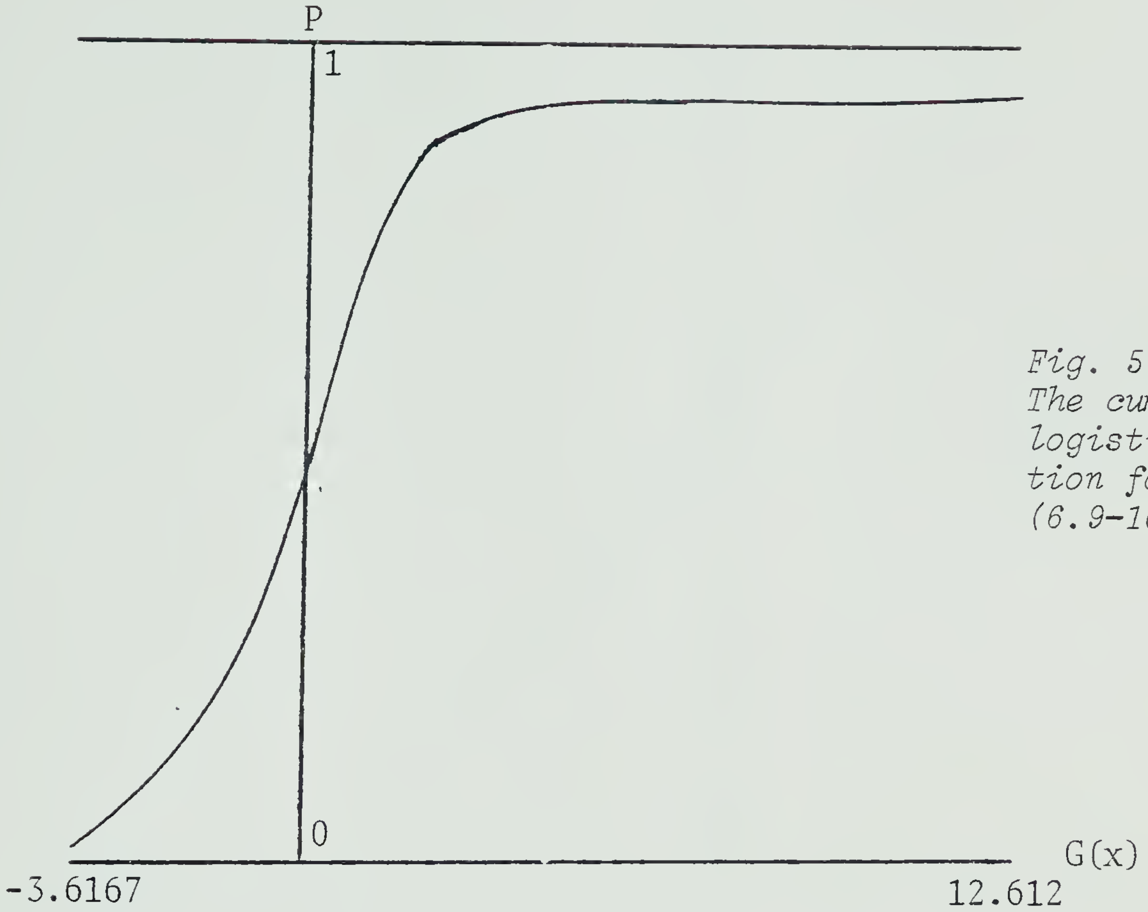


Fig. 5.4
The cumulative logistic function for model 3 (6.9–10.8 mi)

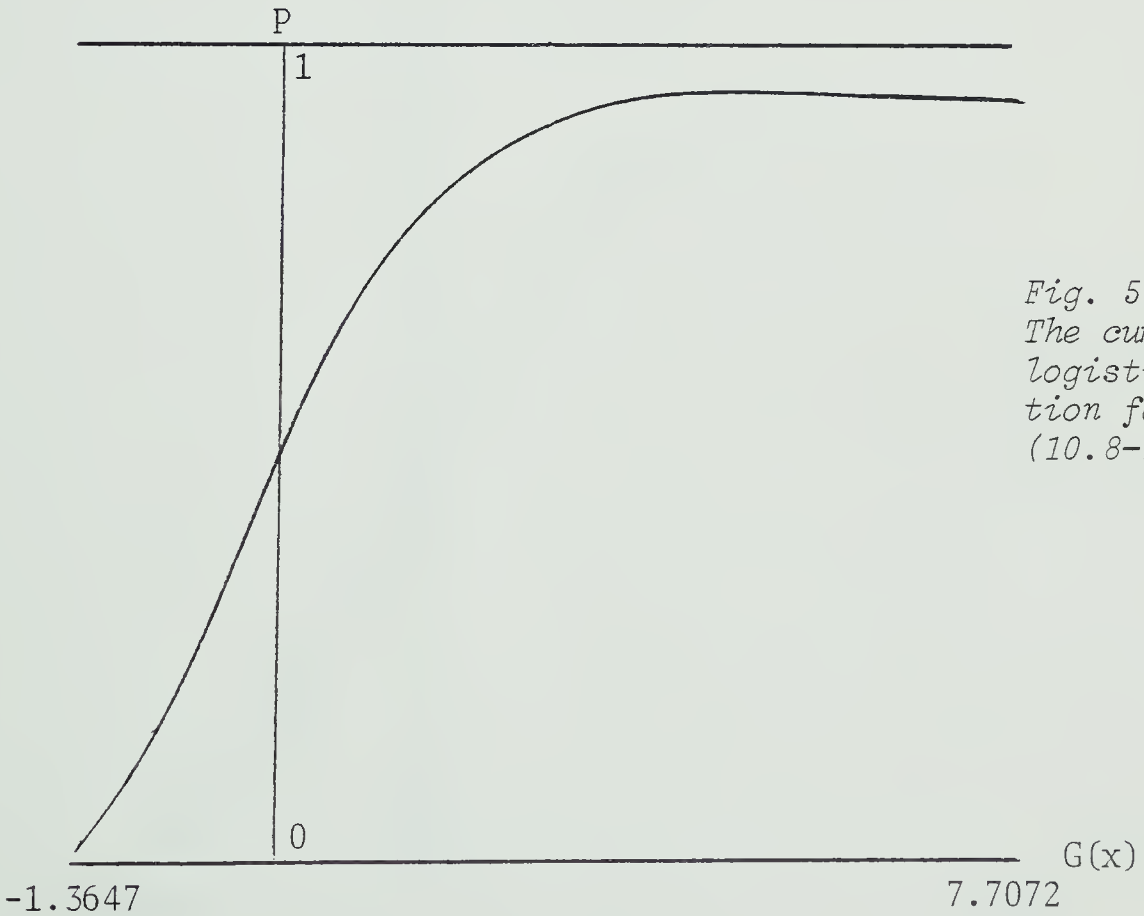


Fig. 5.5
The cumulative logistic function for model 4 (10.8–16 mi)

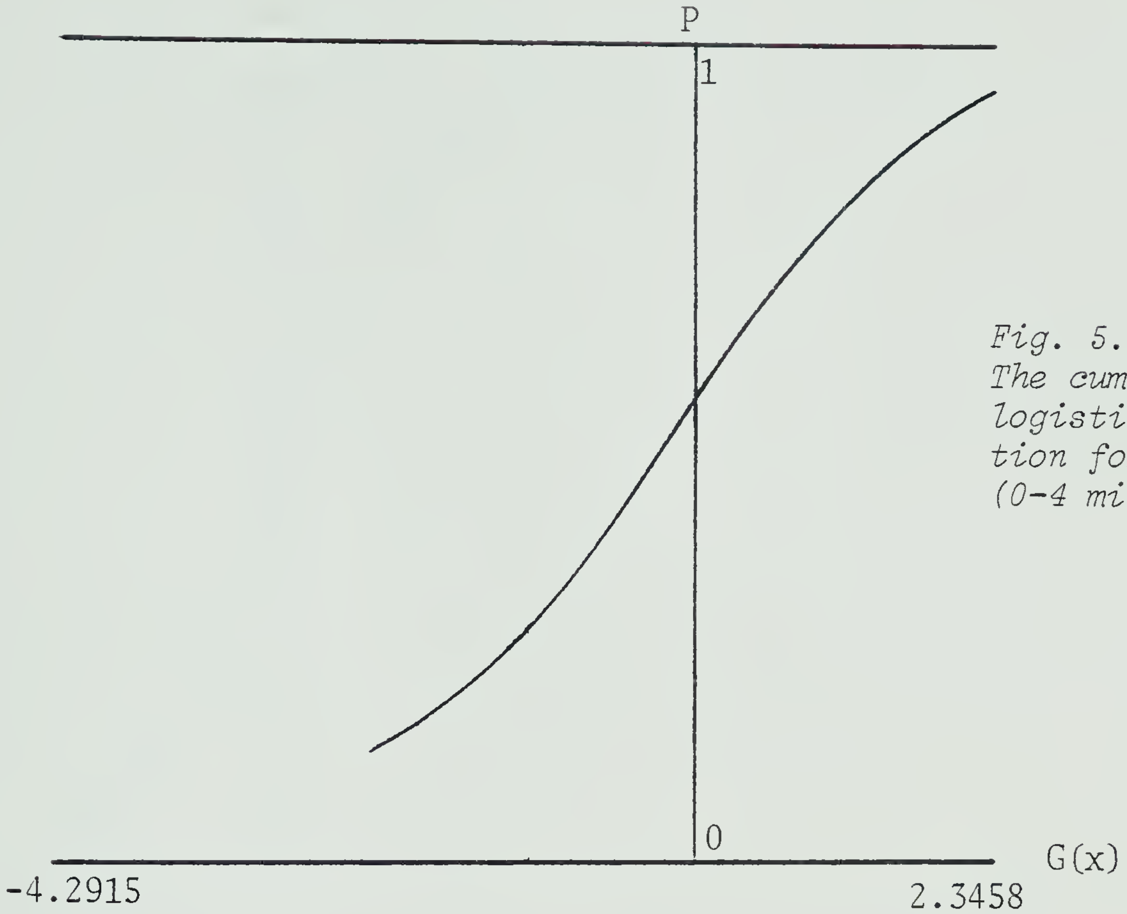


Fig. 5.6
*The cumulative logistic function for model 1**
(0-4 mi)

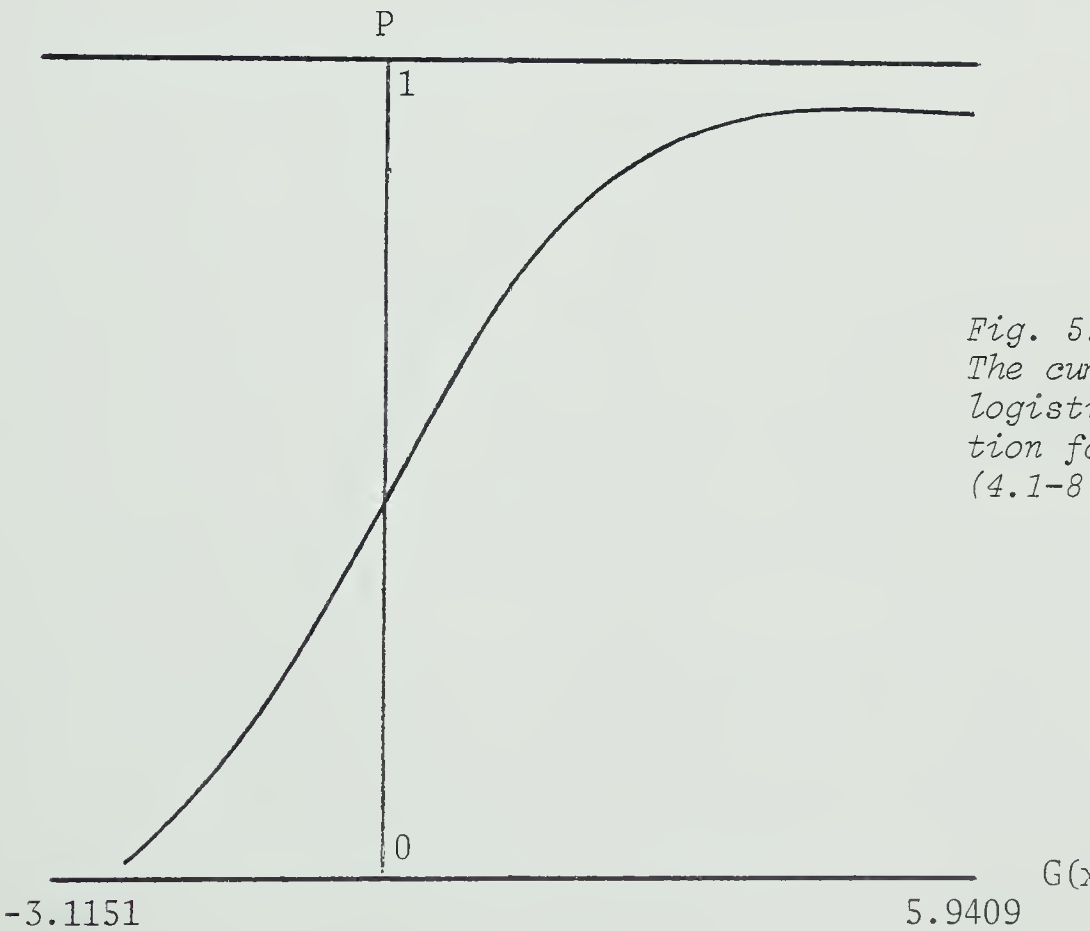


Fig. 5.7
*The cumulative logistic function for model 2**
(4.1-8 mi)

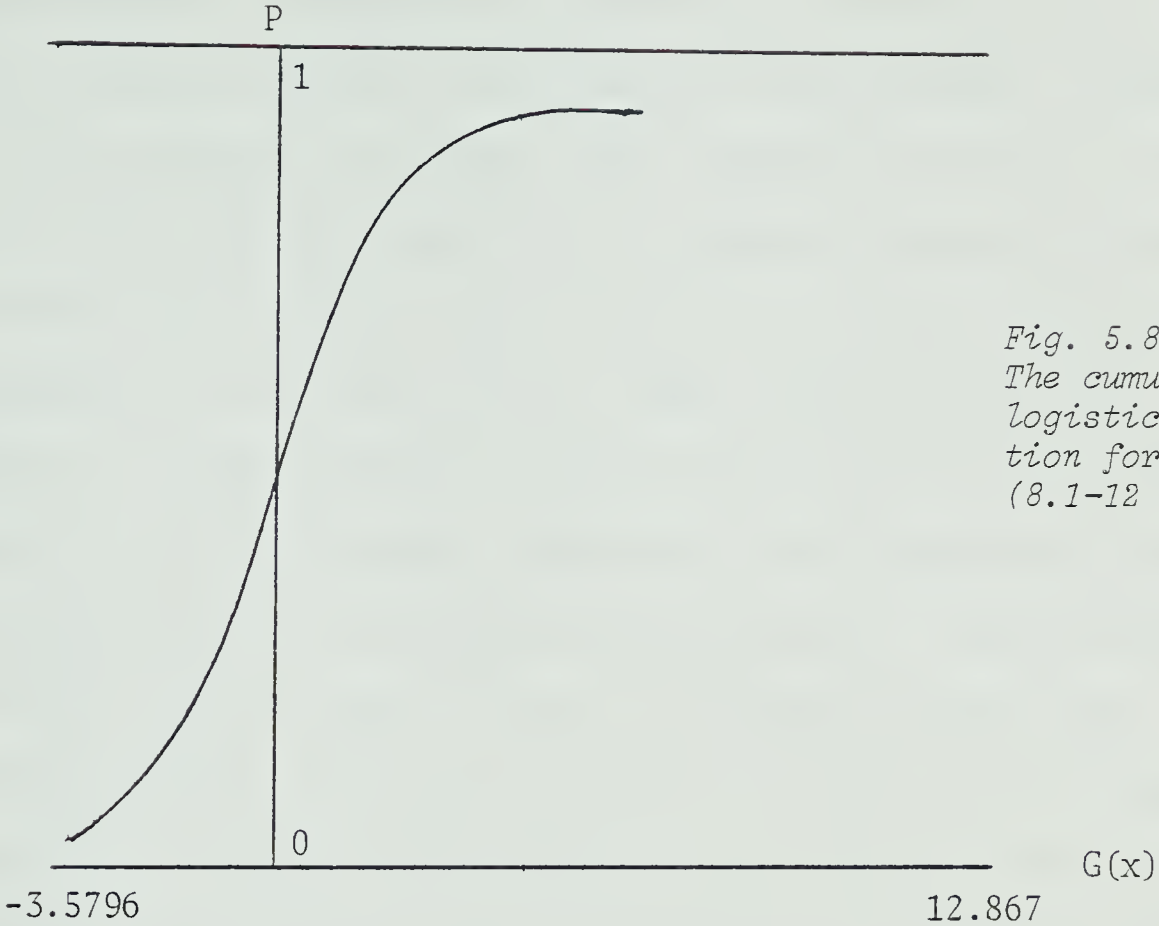


Fig. 5.8
*The cumulative logistic function for model 3**
(8.1-12 mi)

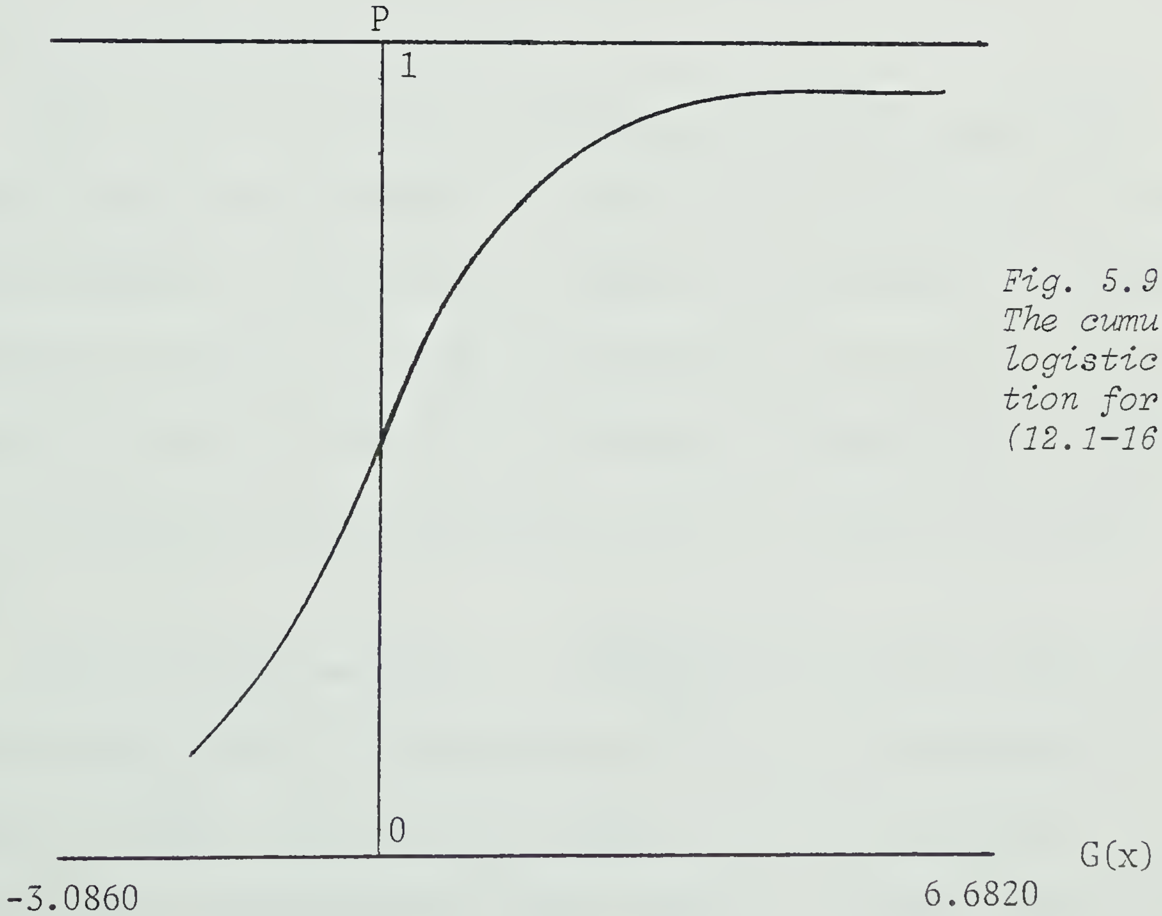


Fig. 5.9
*The cumulative logistic function for model 4**
(12.1-16 mi)

1. From Table 5.9 a comparison of the actual proportion of auto users and the estimated proportion E(P) indicates that the model performs well in the estimation of the model split between the auto and the bus.
2. The results indicate that there is a generally increasing pattern of auto use as one moves away from the city centre, although the proportion of auto users in each sample does not vary widely.
3. There is a wider variation in probabilities near the city centre than at a further distance, indicating a wider variation in observed characteristics of individuals and alternatives. The area models 3 and 3^* show the least variance in individual probabilities.
4. The mean values of the characteristics index G(x) show an expected pattern of increase as one moves away from the city centre. This reflects mainly the increases in time and cost difference facing the individuals.

5.3.2.3 Elasticity calculations

The elasticity formulas derived from Westin's aggregation procedure (described in Chapter 3) describes the sensitivity of changes in the probability distribution to changes in explanatory variables. The elasticity between the expected proportion of auto users E(P) and changes in the mean values of G(x) provides the causal link between policy-related variables and behavior which is required by policy makers.

The elasticity measure, $E = \dfrac{\beta_k \mu_k E[P(1-P)]}{E(P)}$ borrows information from the distribution of probabilities (Table 5.9), parameters of the observed preference of individuals (Table 5.7), and parameters of the characteristics distributions in each sample. The mean values of LOS

TABLE 5.10

ZONAL MEANS FOR SPECIFIED LOS VARIABLES

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
MIN	-17.967	-14.273	-19.192	-18.785	-21.029	-14.387	-17.380	-21.00	-23.984
COST	8.5267	7.4545	7.6946	9.4861	9.9904	6.8763	8.7167	9.8743	8.7705
PCOST	22.374	22.010	22.389	22.677	22.346	22.000	22.473	22.527	22.525

variables are indicated in Table 5.10. This distributional information is combined with information summarizing individual behavior in order to calculate elasticities for each model (see Table 5.11).

Several observations can be made about these elasticities:

1. The elasticities vary as one moves across the city.
2. The time elasticities show an increasing pattern as distance from the city centre is increased. These peak in areas of 3 and 3*, and drop off again for those located at the farthest distance from the city centre. This is likely because of increasing time deviations between modes as one moves away from the city centre.
3. Individuals in these samples are more sensitive to time savings than they are to cost savings. This might reflect the high average incomes of the city.
4. One would expect, therefore, that the areas which have the highest time elasticities would have the lowest cost elasticities. This relationship does not hold, especially for area model 3. This leads one to suspect that other effects might be captured in the time responses for that area.
5. The elasticities with respect to parking costs seem to increase as one moves farther from the city core. Interpretation of the pattern of parking elasticities is difficult, especially due to the lack of more detailed information about parking behavior,[17] and due to the deficiency in the measurement of the parking costs through averaging. Also, since the elasticity is a random variable, composed of behavioral parameters of questionable significance in some base models (see Table 5.7), the significance of the elasticity measure itself is questionable.

TABLE 5.11

ACTUAL ELASTICITIES FOR EACH DISTANCE STRATIFIED MODEL

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.8-16	0-4	4.1-8	8.1-12	12.1-16
Elasticities:									
MIN	.4777	.3643	.3771	.6188	.4900	.3437	.4596	.6384	.3699
COST	-.0358	-.0355	-.0240	-.0484	-.0484	-.0011	-.0581	-.0082	-.0877
PCOST	-.2848	-.0923	-.2918	-.5204	-.3576	-.0273	-.3687	-.8079	-.9233

5.3.2.4 Tests of transferability

If a policy maker does not have access to information from a base model for a specific area in order to calculate elasticities for that area, then it may be possible to borrow this information from other models. The constraints in the use of this procedure depend upon the outcomes of tests for transferability.

One test for transferability is the Likelihood Ratio Test, where the null hypothesis is that all or a subset of the coefficients of the models are equal across distance-stratified models. These tests will be done at three levels.

1. A test of the hypothesis that the entire model is transferable across nine distance mocels.
2. A test of the hypothesis that the parameters of all LOS variables are transferable.
3. A test that the parameter of each separate LOS variable is transferable across the distance models.

The test statistic $-2\ell n\lambda$ is distributed asymptotically chi square with K^* degrees of freedom. As previously indicated, K^* is the number of constraints in the null hypothesis, and λ is the ratio of likelihoods of a restricted and unrestricted model.

The transferability test involves the estimation of the model on a restricted and unrestricted data set for each pair-wise combination of the nine distance models. Suppose for illustration that an entire vector of coefficients is tested between two data sets of equal size. The restricted data set involves K^* parameters constrained in the null hypothesis and is estimated on a combined sample of $2n \times K^*$ observations. The unrestricted sample would have $2n \times 2K^*$ observations. Tests on

subsets of coefficients would involve a restricted sample of $[2n\times(K^{*}+2(K-K^{*})]$ observations, and an unrestricted sample of $2n\times 2K^{*}$ observations.

Fortunately, Westin and Watson [59] state an equivalent and less complicated procedure where the likelihood of the unrestricted sample is the product of the likelihoods of the two separate data sets. This simplifies the procedure considerably.

Tables 5.12 to 5.21 show the calculated test statistics for the two stratification procedures under the question of the transferability of specific sets of variables.

The test results are most clear when illustrated diagrammatically as they show often wide variation.[18] Figures 5.10 to 5.14 indicate the values of the test statistics for each combination of data tested, and critical levels of confidence.

The following observations are made about the test results:

1. The test statistics displayed in Fig. 5.10 where the hypothesis is that the entire equation is transferable indicate that the hypothesis cannot be rejected at any reasonable level of confidence. However, two plausible explanations exist for these results. One is that transferability is strong, and that the coefficients of all LOS and socioeconomic variables are transferable between the samples. This would indicate that the model has identified representative and common portions of utility. The other is that the acceptance of the hypothesis of transferability is due to the liberal degrees of freedom allowed to the test statistics, given that many of the socioeconomic variables are insignificant at reasonable levels of confidence. If the constraint of the equality of the socioeconomic

TABLE 5.12

LIKELIHOOD RATIO TEST STATISTICS FOR LOGARITHMIC MODELS 1-4 WITH SEVEN DEGREES OF FREEDOM

	City (0-16) mi	1 (0-4.3) mi	2 (4.4-6.8) mi	3 (6.9-10.8) mi	4 (10.8-16) mi
City (0-16)	-	2.98	3.16	5.84	0.484
1 (0-4.3)	-	-	4.78	9.66	0.964
2 (4.4-6.8)	-	-	-	7.92	1.544
3 (6.9-10.8)	-	-	-	-	2.704

TABLE 5.13

LIKELIHOOD RATIO TEST STATISTICS FOR EQUAL DISTANCE MODELS 1^*-4^* WITH SEVEN DEGREES OF FREEDOM

	City (0-16) mi	1* (0-4) mi	2* (4.1-8) mi	3* (8.1-12) mi	4* (12.1-16) mi
City (0-16)	-	3.62	1.72	5.614	5.13
1* (0-4	-	-	6.6	8.534	6.13
2* (4.	-	-	-	6.068	3.65
3*	-	-	-	-	8.064

TABLE 5.14

LIKELIHOOD RATIO TEST STATISTICS FOR LOGARITHMIC MODELS 1-4 WITH THREE DEGREES OF FREEDOM

	City (0-16) mi	1 (0-4.3) mi	2 (4.4-6.8) mi	3 (6.9-10.8) mi	4 (10.8-16) mi
City (0-16)	-	2.5	3.12	4.5	2.44
1 (0-4.3)	-	-	4.44	7.24	.744
2 (4.4-6.8)	-	-	-	7.08	1.484
3 (6.9-10.8)	-	-	-	-	2.044

TABLE 5.15

LIKELIHOOD RATIO TEST STATISTICS FOR EQUAL DISTANCE MODELS 1^*-4^* WITH THREE DEGREES OF FREEDOM

	City (0-16) mi	1* (0-4) mi	2* (4.1-8) mi	3* (8.1-12) mi	4* (12.1-16) mi
City (0-16)	-	3.16	1.0	5.414	2.43
1* (0-4)	-	-	4.64	8.3451	3.09
2* (4.1-8)	-	-	-	5.274	1.63
3* (8.1-12)	-	-	-	-	4.864

TABLE 5.16

LIKELIHOOD RATIO TEST STATISTICS FOR LOGARITHMIC MODELS 1-4
WITH ONE DEGREE OF FREEDOM (MIN)

	City (0-16) mi	1 (0-4.3) mi	2 (4.4-6.8) mi	3 (6.9-10.8) mi	4 (10.8-16) mi
City (0-16)	-	2.24	1.78	4.26	0.02
1 (0-4.3)	-	-	2.46	7.06	0.644
2 (4.4-6.8)	-	-	-	5.86	0.444
3 (6.9-10.8)	-	-	-	-	1.940

TABLE 5.17

LIKELIHOOD RATIO TEST STATISTICS FOR EQUAL DISTANCE MODELS 1*-4*
WITH ONE DEGREE OF FREEDOM (MIN)

	City (0-16) mi	1* (0-4) mi	2* (4.1-8) mi	3* (8.1-12) mi	4* (12.1-16) mi
City (0-16)	-	2.6	0.48	4.35	0.93
1* (0-4)	-	-	3.38	8.014	1.29
2* (4.1-8)	-	-	-	3.214	0.61
3* (8.1-12)	-	-	-	-	2.524

TABLE 5.18

LIKELIHOOD RATIO TEST STATISTICS FOR LOGARITHMIC MODELS 1-4
WITH ONE DEGREE OF FREEDOM (COST)

	City (0-16) mi	1 (0-4.3) mi	2 (4.4-6.8) mi	3 (6.9-10.8) mi	4 (10.8-16) mi
City (0-16)	-	0.88	1.14	0.34	0.064
1 (0-4.3)	-	-	0.246	1.18	0.210
2 (4.4-6.8)	-	-	-	0.48	0.404
3 (6.9-10.8)	-	-	-	-	0.064

TABLE 5.19

LIKELIHOOD RATIO TEST STATISTICS FOR EQUAL DISTANCE MODELS 1^*-4^*
WITH ONE DEGREE OF FREEDOM (COST)

	City (0-16) mi	1^* (0-4) mi	2^* (4.1-8) mi	3^* (8.1-12) mi	4^* (12.1-16) mi
City (0-16)	-	1.02	0.82	0.954	1.95
1^* (0-4)	-	-	2.42	0.854	2.91
2^* (4.1-8)	-	-	-	1.534	0.99
3^* (8.1-12)	-	-	-	-	2.244

TABLE 5.20

LIKELIHOOD RATIO TEST STATISTICS FOR LOGARITHMIC MODELS 1-4 WITH ONE DEGREE OF FREEDOM (PCOST)

	City (0-16) mi	1 (0-4.3) mi	2 (4.4-6.8) mi	3 (6.9-10.8) mi	4 (10.8-16) mi
City (0-16)	-	0.88	0.292	0.80	0.124
1 (0-4.3)	-	-	3.7	1.62	0.204
2 (4.4-6.8)	-	-	-	2.02	1.184
3 (6.9-10.8)	-	-	-	-	0.064

TABLE 5.21

LIKELIHOOD RATIO TEST STATISTICS FOR EQUAL DISTANCE MODELS 1^*-4^* WITH ONE DEGREE OF FREEDOM (PCOST)

	City (0-16) mi	1* (0-4) mi	2* (4.1-8) mi	3* (8.1-12) mi	4* (12.1-16) mi
City (0-16)	-	0.8	0.64	1.994	0.75
1* (0-4)	-	-	0.16	3.314	1.19
2* (4.1-8)	-	-	-	2.179	0.53
3* (8.1-12)	-	-	-	-	0.504

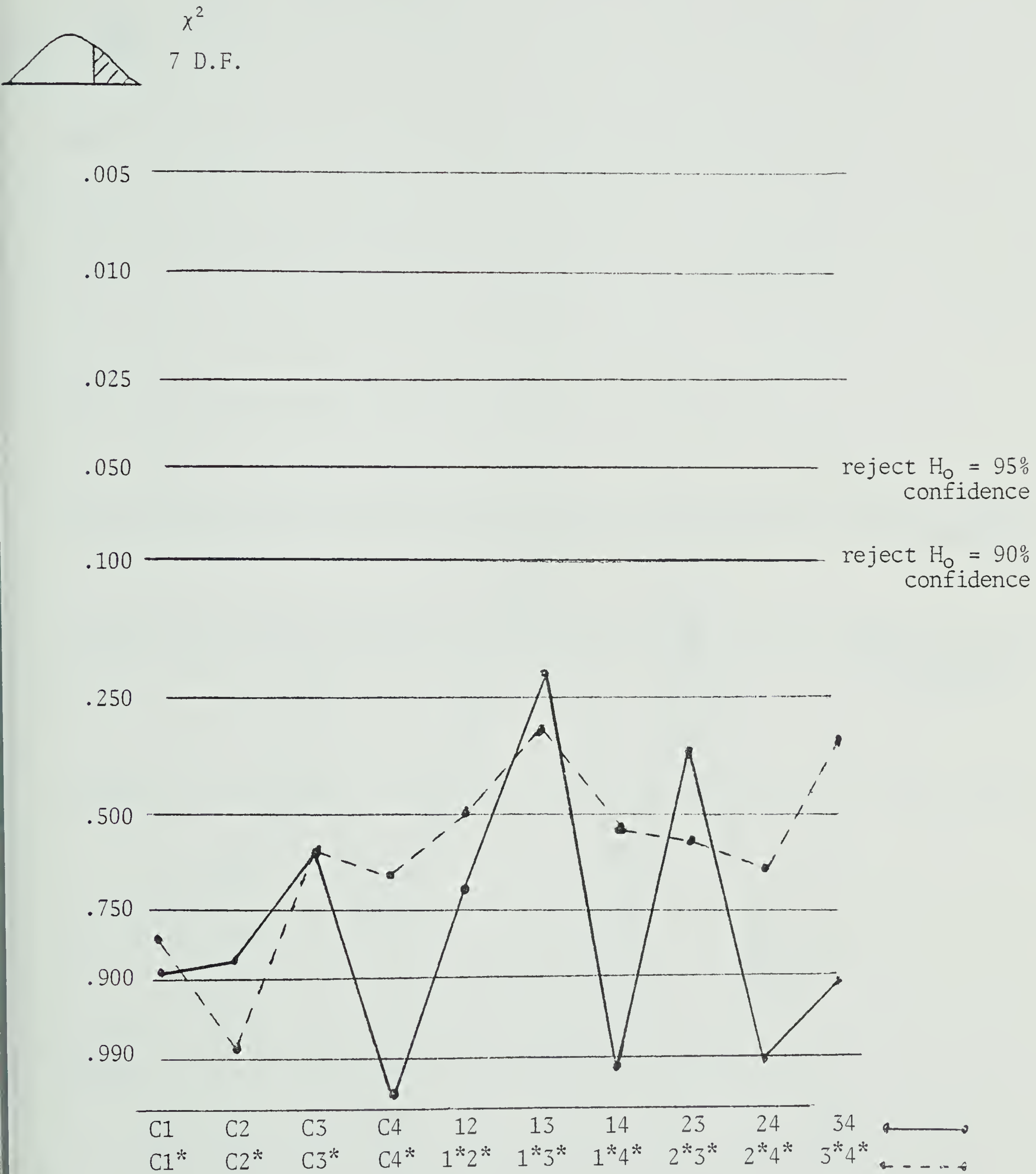


Fig. 5.10 Likelihood ratio test statistics where H_o = equality of coefficients of entire equation across all samples.

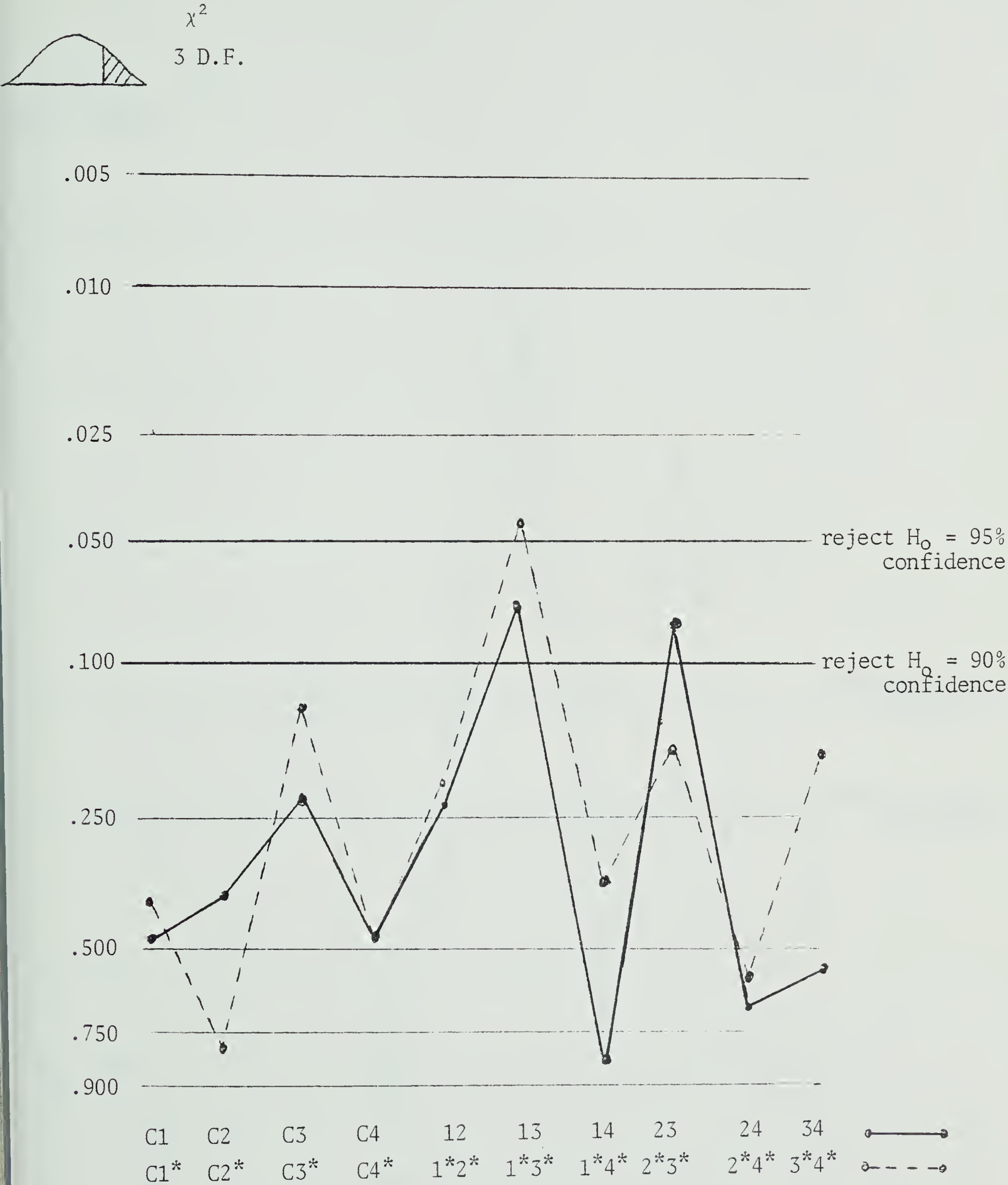


Fig. 5.11 *Likelihood ratio test statistics where H_o = equality of coefficients of LOS variables across all samples.*

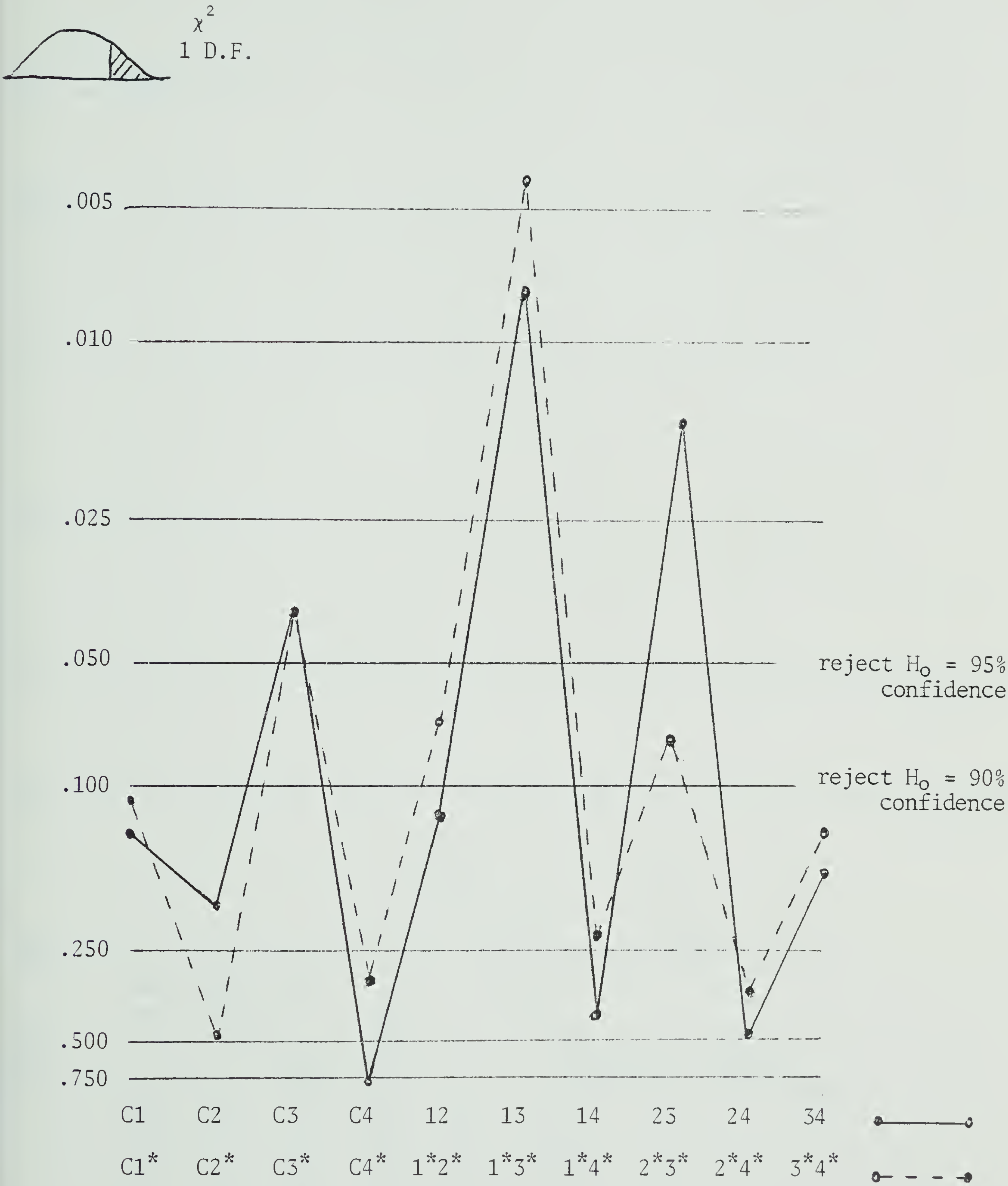


Fig. 5.12 Likelihood ratio test statistics where H_o = equality of MIN coefficients across all samples.

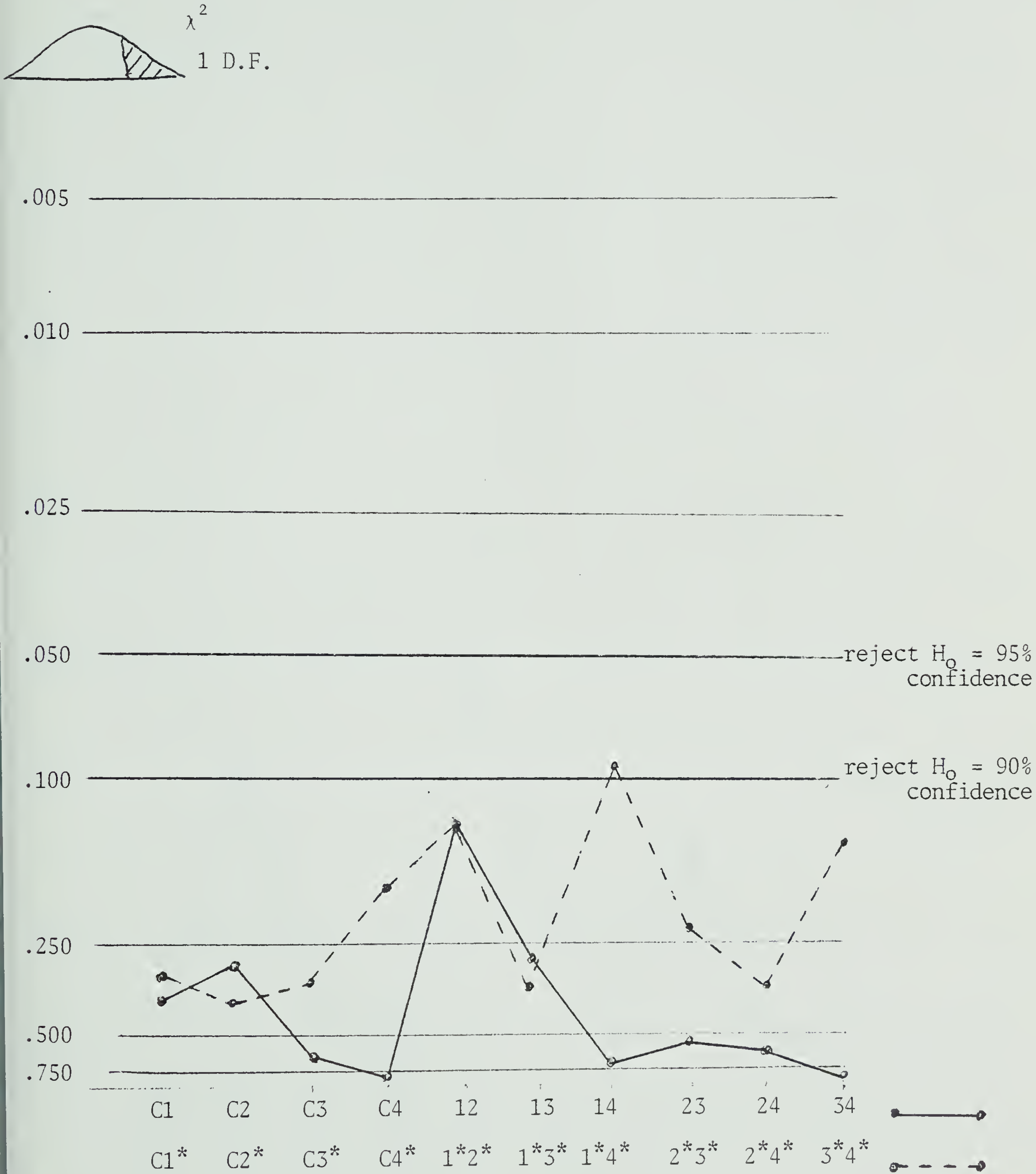


Fig. 5.13 Likelihood ratio test statistics where H_o = equality of COST coefficients across all samples.

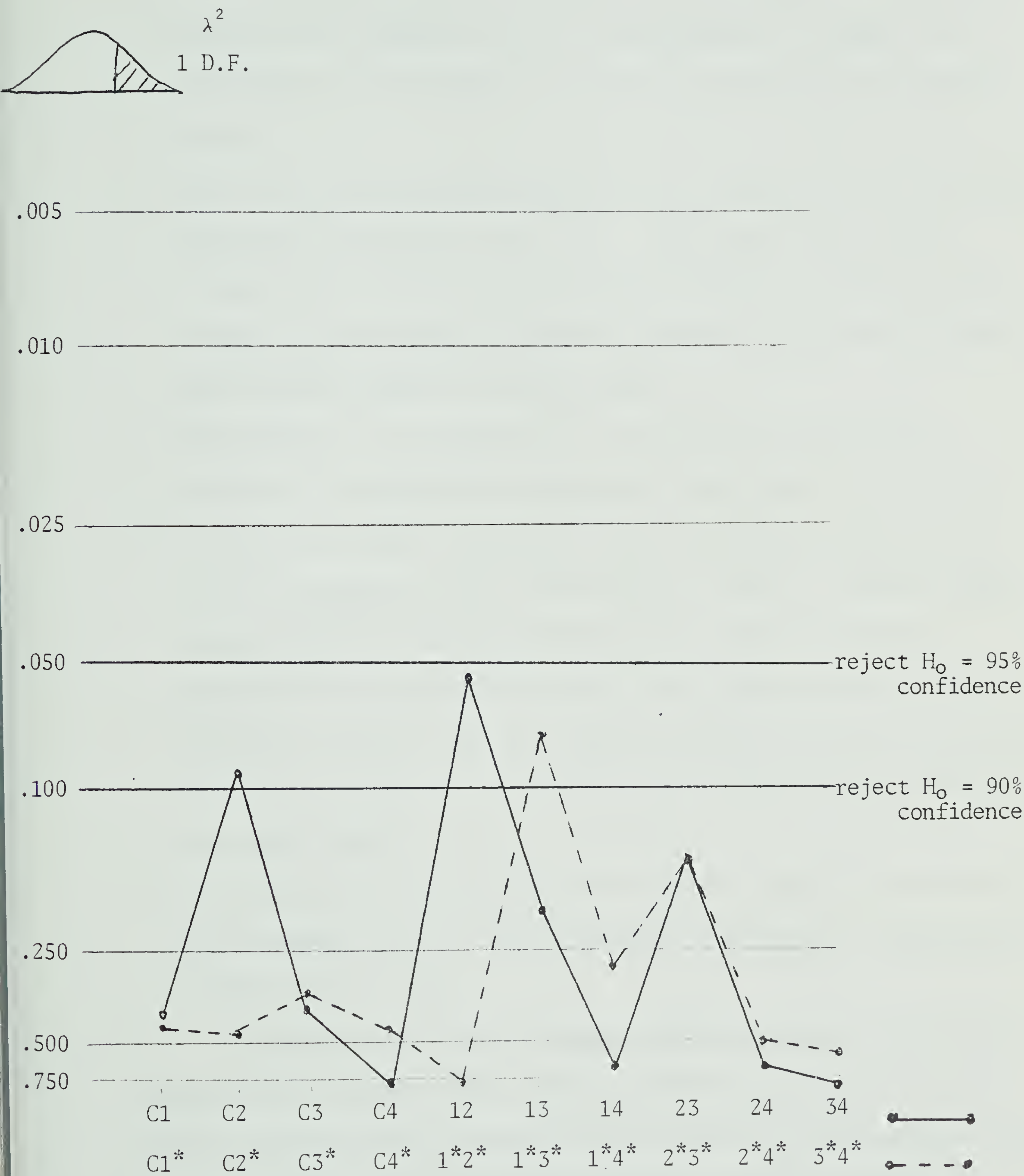


Fig. 5.14 Likelihood ratio test statistics where H_o = equality of PCOST coefficients across all samples.

variables was relaxed, then in this case one might observe almost identical test statistics but at lower degrees of freedom. This would make it more difficult to accept the hypothesis of transferability.

2. The second explanation above seems to be confirmed when one examines the results displayed in Fig. 5.11. The values of the highest original test statistic have decreased slightly, but remain high enough at three degrees of freedom to reject one case at 95% confidence and two others at the 90% level.
3. The hypothesis of transferability has not been systematically rejected by the majority of subsample comparisons.
4. In almost all cases, the highest value of the test statistics for the LOS variables are for distance models 3 and 3^*. However, good transferability is indicated between quite spatially diverse populations such as between models 1 and 4, and 1^* and 4^*. This initially indicates that there may be unobserved effects capitalized in the LOS variables in the areas (3) (3^*) which are not present in the other samples.
5. The pattern of results becomes less dependent upon the stratification procedure, as the less significant socioeconomic variables are removed from the test.

A similar set of tests on single coefficients, on a variable-by-variable basis between samples, serves several purposes. One is to further pursue the results above--in this case being the suggestion that the less significant pcost and cost variables influence the test in making it easier to accept the hypothesis of transferability. The second reason is to examine whether each of the LOS variables test

results are of a similar pattern, and are as independent of the stratification procedure as seems to be indicated in Fig. 5.11.

By examining the results in Figs. 5.12 to 5.14 we observe that:

1. The time coefficients show a strong pattern of transferability success which is seemingly independent of the stratification procedure.
2. The time results show mixed success between areas, with transferability being strongly rejected between areas 3 and 3* and the other models, but generally being accepted between the other regions.
3. The cost and parking coefficients seem to be generally transferable with two exceptions in pcost, but with a less obvious pattern. The test results are much more dependent upon the sample selection than are the time coefficients.
4. By comparing the similarity of Figs. 5.11 and 5.12, it seems that the general LOS test was weighted by the strong results of the time test, and that similarity of the results in Fig. 5.11 is more likely due to this effect than was due solely to the removal of socioeconomic variables.

The special pattern of results for areas 3 and 3* warrant further discussion.

Theoretically, transferability failure is due to nonindependence between unobservable and observable characteristics. Failure to transfer the coefficients, in this case, must therefore be due to the time coefficient capitalizing some effects not captured in the model, but correlated to time differences. In practice, this could be due to inadequate model specification as well as nonindependence of unobservable

and representative preferences. If the model was generally mis-specified, however, one would expect a larger mix of transferability success and failure, involving more variables, in more areas.

These results must be explained in terms of a model mis-specification affecting the time difference variable for areas 3 and 3*.

Examination of Table 5.22 indicates that some individuals in areas 3 and 3* seem to be disadvantaged with respect to some time and cost variables. This is not an unlikely result, as these may represent areas which are quite far from the city centre, but may not be far enough to have received the benefits of better highways or special express bus service, such as one might expect for distance areas 4 and 4*.

TABLE 5.22

COMPARISON OF REPORTED RANGES FOR TIME AND COST
IN OUTER DISTANCE MODELS

BETWEEN AREAS 3, 4					BETWEEN AREAS 3*, 4*			
		(Miles)	Low	High		(Miles)	Low	High
auto t	(3)	(6.9-10.8)	10	60	(3*)	(8.1-12)	10	60
	(4)	(10.8-16)	10	60	(4*)	(12.1-16)	12	60
bus t	(3)	(6.9-10.8)	15	150	(3*)	(8.1-12)	20	150
	(4)	(10.8-16)	15	120	(4*)	(12.1-16)	15	120
auto c	(3)	(6.9-10.8)	2	100	(3*)	(8.1-12)	4	120
	(4)	(10.8-16)	2	120	(4*)	(12.1-16)	2	70
bus c	(3)	(6.9-10.8)	4	90	(3*)	(8.1-12)	4	90
	(4)	(10.8-16)	5	66	(4*)	(12.1-16)	5	66

It could be possible that the transit planners compete with the auto in these areas (where they are constrained from doing so by time or cost reduction) through the offering of special services to counteract the disadvantages. Such aspects as comfort or convenience, for example, could be closely correlated with time costs. We lack adequate proxies for this analysis for these other variables which may then systematically differ across different parts of the city.

In general, in testing for transferability across space, the variability in service levels between particular areas (and when the service levels are in difficult to measure or are unobservable effects) may cause nontransferability of results. It is probable that these service levels will be correlated very closely to distance, with the distance disadvantage for transit being reflected in travel time. (Fares usually do not differ significantly.)

Transit authorities will have to compensate for time disadvantages of distance in increasingly large urban centres in order to preserve ridership. This would be reduced in cities where the auto does not have such a large time advantage over public transit (i.e., where high-speed public transit modes are available, or in smaller urban areas where distance differences are less pronounced).

Examination of the reported minimum and maximum cost differences for areas in these models seem to support this hypothesis. These are contained in Table 5.23. It is therefore suggested that the variability in travel times between modes is one of the relevant variables to be examined in the a priori expectations of transferability success.

In this analysis there does not seem to be enough systematic rejection of transferability of all coefficients to reject the notion

of transferability across space per se.

TABLE 5.23

MAXIMUM AND MINIMUM REPORTED MIN FOR AREAS 3,4 and 3*,4*

T_A-T_C	Low	High
3	-115.00	10
4	-100.00	10
3*	-115.00	10
4*	-100.00	0

5.3.2.5 Predictions using transferred parameters

The ultimate benefit from tests of transferability will be in the implications for policy makers. At this point we have merely established that we cannot reject totally the idea of spatial transferability of coefficients with the Ottawa/Hull area, and that the case of strong rejection observed might have been predicted. It would be useful also to compare the prediction results for populations having borrowed parameters from other populations. Such predictions would be in the estimation of modal splits and elasticities generated from parameters of different models. Given that there is evidence to support transferability in general, a few options are available to the policy maker to make use of the information.

Suppose that the city planner has an estimated city model (or other subarea model) and distributional information on subsamples of his population. One way of taking into account the fact that elasticities will differ across urban space is to use only the new distributional information and calculate new elasticities. In this analysis, using only the mean values of characteristics as new information, and using a base city model, we get the following results (Table 5.24).

TABLE 5.24

COMPARISON OF ACTUAL ELASTICITIES AND THOSE CORRECTED ONLY FOR DIFFERENCES IN CHARACTERISTICS DISTRIBUTIONS

MODEL	CITY	1	2	3	4	1*	2*	3*	4*
Distance range (mi)	0-16	0-4.3	4.4-6.8	6.9-10.8	10.9-16	0-4	4.1-8	8.1-12	12.1-16
Min		.3759	.5103*	.4995	.5591	.3825	.4621	.55841	.6377
(Original *E*)	(.4777)	(.3643)	(.3771)	(.6188)	(.4900)	(.3437)	(.4596)	(.6384)	(.36998)
Cost		-.0313*	-.0323	-.0399	-.04198	-.02889	-.0366	-.0418	-.03619
(Original *E*)	(-.0358)	(-.0355)	(-.0240)	(-.0484)	(-.0484)	(-.0011)	(-.0581)	(-.00824)	(-.0877)
pcost		-.2801	-.28495	-.2886	-.2844	-.2800	-.2860	-.2867	-.2867*
(Original *E*)	(-.2848)	(-.0923)	(-.2918)	(-.5204)	(-.3576)	(-.0273)	(-.3687)	(-.80796)	(-.9233)

*indicates that city-wide elasticity is closer to the actual elasticity than those corrected for only changes in characteristics.

It can be seen that the general pattern of the elasticities is preserved, although of course the range of elasticities is considerably reduced. In all cases except those marked * above, if the original elasticity is considered to be the "true value", then these new elasticities are closer to the true values than is the city-wide elasticity. This implies that the distributions of characteristics are important to elasticity calculations. However, all the elasticities above are random variables, and this pattern may not hold in other cases. This rather naive methodology of course imposes an assumption of transferability of *all* of the information required in the elasticity except specific distributional statistics.

Despite this rather restrictive assumption, in this case above, the policy maker has estimates of elasticities more closely resembling the "true" calculated values than if this methodology had not been applied.

Although the new elasticities are closer to the real values than the city-wide elasticities, there remains a large amount of difference in their magnitudes. This implies that a potentially more accurate method to calculate these elasticities (without a new model) would be to generate new distributional information using the city-wide coefficients, and then to generate the elasticities. Westin's S_B distribution is such a generated distribution. This procedure requires exactly the same input as one would already have if one were to calculate the elasticities using the immediately preceding approach.

The calculation of the S_B distribution is described in Chapter 3, which has a probability density function equal to (3.6).

$$f(p) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{P_j^a(1-P_j^a)} \exp\left\{\frac{-1}{2\sigma^2} \left| \ell n(\frac{P}{1-P}) - \mu \right|^2 \right\}$$

This distribution has parameters (μ,σ) which are linear combinations of the information which the city planner already has from the city base model, and the distributional information of the subareas, i.e., $\mu = \mu_k \beta_k$ and $\sigma = \beta\Sigma\beta'$ as previously described in Chapter 3.

The descriptive parameters (μ,σ) to be used in the generation of the S_B distribution are contained in Table 5.25. The S_B distribution can be generated with numerical integration procedures. The summary statistics for the distributions which will be used in the calculation of new elasticities (and E(P) for its interpretation as an expected proportion of travellers of taking the auto) are recorded in Table 5.26. The models C1 to C4* once again represent subarea models which have the S_B distribution calculated from coefficients of the city-wide model. Potentially, all subarea models have generated probability distributions using each other area. The generation of the S_B distribution is, however, a complicated procedure, and it is felt that the use of the city-wide model will suffice for purposes of illustration. Also, the city-wide model will most usually be the model that is available to policy makers, and would be used to generate information about its subareas when necessary.

We can determine the success of Westin's procedure by comparing the distribution moments E(P) for each subarea in Table 5.27 with the corresponding moments from the actual subarea probability distributions.

TABLE 5.25

GENERATED PARAMETERS (μ,σ) OF THE S_B DISTRIBUTION

MODEL	DISTANCE RANGE	$\mu = (\beta_k \mu_k)$	$\sigma = \beta\Sigma\beta'$	DISTANCE MODELS WITH CITY COEFFICIENTS	$\mu = (\beta_k \mu_k)$	$\sigma = \beta\Sigma\beta'$
CITY	0-16	.61749	1.3383	-	-	-
1	0-4.3	.26358	.8987	C1	.30525	1.0886
2	4.4-6.8	.75539	1.0696	C2	.74658	1.3192
3	6.9-10.8	.78859	2.1153	C3	.66376	1.4341
4	10.9-16	.93890	1.4798	C4	.87795	1.4880
1*	0-4	.28720	.8108	C1*	.34078	1.1092
2*	4.1-8	.58166	1.3100	C2*	.55589	1.2384
3*	8.1-12	1.0649	2.2075	C3*	.86033	1.5232
4*	12.1-16	1.1645	1.6267	C4*	1.1472	1.7142

TABLE 5.26

SUMMARY STATISTICS OF GENERATED PROBABILITY DISTRIBUTION USING OWN AND TRANSFERRED COEFFICIENTS

MODEL	DISTANCE	$E(P)_{S_B}$	$E(P[1-P])_{S_B}$	GENERATED MODELS	$E(P)'_{S_B}$	$E(P[1-P])'_{S_B}$
CITY	0-16	.61315	.17782			
1	0-4.3	.55591	.21055	C1	.56112	.19847
2	4.4-6.8	.64875	.18646	C2	.63666	.17508
3	6.9-10.8	.61374	.14012	C3	.61786	.17191
4	10.9-16	.66191	.16196	C4	.65161	.16351
1*	0-4	.56229	.21486	C1*	.56773	.19663
2*	4.1-8	.60769	.18018	C2*	.60532	.18458
3*	8.1-12	.64797	.13218	C3*	.64718	.16255
4*	12.1-16	.68951	.14913	C4*	.68232	.14674

TABLE 5.27

COMPARISON OF ACTUAL E(P) AND THOSE GENERATED FROM THE S_B DISTRIBUTION

MODEL	DISTANCE RANGE	E(P) ACTUAL (± ERROR = 1.96 DEV.)	$E(P)_{S_B}$	$E(P)'_{S_B}$ USING CITY COEFFICIENTS (MODELS C1-C4*)
CITY	0-16	.6066 (±.0298)	.6132	-
1	0-4.3	.5522 (±.0622)	.5559	.5611
2	4.4-6.8	.6358 (±.0594)	.6488	.6367
3	6.9-10.8	.5976 (±.0463)	.6137	.6179
4	10.9-16	.6539 (±.0774)	.6619	.6516
1*	0-4	.5645 (±.0666)	.5623	.5678
2*	4.1-8	.6006 (±.0443)	.6077	.6053
3*	8.1-12	.6407 (±.0551)	.6479	.6471
4*	12.1-16	.6885 (±.0969)	.6895	.6823

The moment E(P) is chosen for comparison because of its policy implications. We are also able to compare the quality of predictions with transferable models by comparing these moments generated through the use of the coefficients of city-wide models. Both of these comparisons can be made by observing the results contained in Table 5.27.

We see that the aggregation procedure itself performs well. The $E(P)_{S_B}$ are within the one percent actual E(P).[19] (The E(P) of even the actual probability is itself a random variable as has been estimated from each subarea model.) This is also true for $E(P)_{S_B}'$ which are the models generated using the city-wide coefficients. This also seems to hold for areas 3 and 3*, where transferability of the coefficients did not hold, which implies that Westin's criteria of equality of the coefficients may be too strict.

Elasticities can be calculated using the moments of the generated probability distributions in the same manner as the elasticities of Table 5.11 were generated from moments of the actual probability distributions. These elasticities are contained in Table 5.28.

The elasticity formula involves the use of several different random variables, β_k, E(P), E(P[1-P]) so that the elasticity itself is a random variable. Theoretically then there should be confidence limits associated with each original elasticity, so that a comparison could be made between the original elasticity and its allowable error and the generated elasticity. Because of the number of random variables in each elasticity measure, the error allowance for each is not obvious and has not been calculated here. However, casual comparisons indicate that the elasticity measures are within reasonable limits of the

TABLE 5.28

COMPARISON OF ACTUAL ELASTICITIES AND THOSE GENERATED FROM AN S_B DISTRIBUTION WITH CITY COEFFICIENTS

MODEL DISTANCE	CITY 0-16	1 0-4.3	2 4.4-6.8	3 6.9-10.8	4 10.9-16	1* 0-4	2* 4.1-8	3* 8.1-12	4 12.1-16
MIN (Original *E*)	 (.4777)	.4247 (.3643)	.4440 (.3771)	.4398 (.6188)	.4439 (.4900)	.4193 (.3437)	.4459 (.4596)	.4438 (.6384)	.4339 (.3699)
COST (Original *E*)	 (-.0358)	-.0351 (-.0355)	-.0281 (-.0240)	-.0351 (-.0484)	-.0333 (-.0484)	-.0317 (-.0011)	-.0353 (-.0581)	-.0329 (-.0082)	-.0251 (-.0877)
PCOST (Original *E*)	 (-.2848)	-.3187 (-.0923)	-.2438 (-.2918)	-.1499 (-.5204)	-.2258 (-.3576)	-.3069 (-.0273)	-.2759 (-.3687)	-.2277 (-.8079)	-.1951 (-.9233)

TABLE 5.29

COMPARISON OF ACTUAL, S_B, AND 'CHARACTERISTICS ADJUSTED' ELASTICITIES

MODEL DISTANCE	1 0-4.3	2 4.4-6.8	3 6.9-10.8	4 10.9-16	1* 0-4	2* 4.1-8	3* 8.1-12	4* 12.1-16
a) Original *E* (MIN)	.3643	.3771	.6188	.4900	.3437	.4596	.6384	.3699
b) Correction for only characteristics differences	.3759	.5103	.4995	.5591	.3825	.4621	.5584	.6377
c) Calculated with generated S_B distribution with city coefficients	.4247	.4440	.4398	.4439	.4193	.4459	.4438	.4339
a) (COST)	-.0355	-.0240	-.0484	-.0484	-.0011	-.0581	-.0082	-.0877
b)	-.0313	-.1323	-.1399	-.0419	-.0289	-.0366	-.0418	-.0362
c)	-.0351	-.0281	-.0351	-.0333	-.0317	-.0353	-.0329	-.0251
a) (PCOST)	-.0923	-.2918	-.5204	-.3576	-.0273	-.3687	-.8079	-.9233
b)	-.2801	-.2850	-.2886	-.2844	-.2800	-.2860	-.2867	-.2867
c)	-.3187	-.2438	-.1499	-.2258	-.3069	-.2759	-.2277	-.1951

originals except in those areas where the coefficients were not transferable. However, due to the similarity of model splits in all areas, and due to the transferability of the coefficients, the elasticities which were corrected by only changing characteristics parameters perform as well as the elasticities operated from the more complicated S_B procedure. This can be seen in Table 5.29. One cannot expect this to hold in all urban areas.

The new generated elasticities do not exactly replicate the actual elasticities. The Likelihood Ratio test for transferability only tests that the *model performs as well* if the coefficients are constrained to take on certain values. It does not determine that the actual point estimates are equal. As long as the point estimates of the coefficients are even slightly different, the actual elasticities cannot be exactly replicated.

5.3.2.6 The aggregate elasticity

One final set of calculations to be provided in this empirical application of the model is with respect to the adequacy of the city-wide model in capturing the overall city response to policy. It is apparent that the subareas will differ in their responses to policy. One wonders if a single city-wide model adequately captures the overall effect of these differential responses.

Domencich and McFadden [9] describe an aggregate elasticity for application to models where market segments are well defined.[20] The aggregate elasticity is defined as being equal to the sum of the segment elasticities, where these are weighted by that segment's share of the total population. If the city-wide model performs well in capturing

TABLE 5.30

COMPARISON OF CITY-WIDE ELASTICITIES AND THE SUM OF SUBAREA RESPONSES

MODEL	DISTANCE RANGE	% SHARE OF TOTAL POPULATION	INDIVIDUAL SUBAREA ELASTICITIES		
			MIN	COST	PCOST
1	0-4.3	27.2	.3643	-.0355	-.0923
2	4.4-6.8	26.5	.3771	-.0240	-.2918
3	6.9-10.8	32.8	.6188	-.0484	-.5204
4	10.9-16	13.6	.4900	-.0484	-.3576
1*	0-4	24.3	.3457	-.0011	-.0273
2*	4.1-8	46.0	.4596	-.0581	-.3687
3*	8.1-12	21.8	.6384	-.0082	-.8079
4*	12.1-16	8.0	.3699	-.0877	-.9233
CITY	0-16	100.0	.4777	-.0358	-.2848
Aggregate Elasticity Models 1-4			.4687	-.0385	-.3217
Aggregate Elasticity Models 1*-4*			.4637	-.0358	-.4262

the overall response, these measures should be similar. Table 5.30 contains the information which would be required for such a calculation and the calculated aggregate elasticities.

Analysis of these results indicates that elasticities calculated from the city-wide model are reasonably similar to the aggregate response from the subareas.

The time elasticities are slightly higher in the city-wide model, i.e., the city-wide model overpredicts response by 1% for a 100% change in time difference, which is a small error. This includes the effects of the overstated responses to areas 3 and 3* which could be understated in the aggregate elasticity because of the arbitrary method that was used to determine the subareas.[21]

The cost elasticities are almost identical under both stratification procedures, indicating that the city-wide model has adequately captured the aggregate response.

The city-wide model's pcost elasticities are the least accurate relative to the sum of the aggregate responses predicted by the area models. The city-wide model underpredicts response by approximately 3.6% for a 100% change in parking costs (models 1-4) to 14% for a 100% change in parking costs (models 1*-4*). However, one is reminded that these elasticities are of questionable validity as they are composed of some parameters which do not have statistical significance.

Chapter 5

FOOTNOTES

[1]The data was intended for use in the assessment of the impact of a change in the Federal Government Parking Policy on employee travel behavior. I am indebted to Dr. D. W. Gillen for allowing me access to this information.

[2]De Leuw Cather Ltd. [8], pp. 4, 75-86.

[3]There were no variables where the sample was entirely complete, so the main ones selected were mode, cost, time, and income.

[4]In the original data, mode choice information was available for four seasons ("this summer", "last winter", "next summer", and "next winter"). However, individuals were asked to respond to other questions with respect to the current time period up to June 1975. The only other information given which varied seasonally was "parking cost" and "walking time". This reflects the original parking survey purpose of the sample. The model for "this summer" was selected for use in this study because the information was most complete for this time period, and because that season most closely related to the reporting period of socioeconomic and LOS variables.

[5]"Workplace" and "home zone" are coded within the sample, so it should have been a simple procedure to eliminate residents reported as being in zones known to be out of the city. Distances up to 65.5 miles away from work are reported. However, the existing description [?] of the sample is deficient in the provision of a key to the coding procedures for the zones, or in the provision of a map indicating the zones. This necessitated broad distance assumptions identifying a Metro Ottawa location.

[6]This would include such inconsistencies as (999)×5¢ bus cost at a distance from work of 1.6 miles. Only very inconsistent reports were deleted, rather than a general deletion of individuals reporting travel time and costs greater than a certain deviation from mean responses. Wide variations in reported time and costs between individuals are observable in this sample. Since the sample is based on perceived values, such a general procedure might have eliminated individuals who select a mode and then adjust their perceptions about alternate modes in order to justify their choice. This could also have deleted individuals from the sample who lack information about alternate modes. A model was tested on a data set including such deletions, and it was confirmed that through the elimination of individuals with large or very small estimates of time and costs, the remaining observations were

clustered in a very narrow distribution. This was especially true with respect to costs. This resulted in the estimation of a cumulative probability function which was likely steeper than is the actual case, and the model performance was poor ($R^2 = 0.08$, Likelihood Ratio Test - 56 with six Degrees of Freedom at a sample size of 656).

[7]All the distance reports were graphed on a logarithmic scale, and then divided into four equal groups. This grouping actually began at a distance of 1.6 miles from the workplace, despite two individuals reporting 0 distance to work. It is apparent that there is a minimum distance from the workplace--where home location is not feasible.

[8]A regression was run between distance and income in order to ensure that these were really not just implicit income stratifications. A partial correlation coefficient equal to 0.09 indicated that this was not the case.

[9]This would be consistent with the deletion of captive bus users within the city.

[10]De Leuw Cather [8], p. 125. The report does not include discussion of the method by which "calculated" values were generated except to say that they were "map measured values". If it is assumed that the techniques were valid, these results would indicate that elasticities calculated from the models using perceived information will make the bus users more elastic with respect to bus time and more inelastic with respect to auto time than would results based on calculated values.

[11]De Leuw Cather [8], p. 128. This would imply that auto drivers would be more elastic with respect to auto cost charges, while bus users would be more inelastic.

[12]Refer to Theil [53], pp. 392-397, and D. McFadden [29], pp. 119-130 and 134-140, for relevant proofs of uniqueness, consistency and normality of estimators, and also for a demonstration that $-2\ell n\lambda$ (where λ is the Likelihood Ratio for the model) has asymptotically a Chi square distribution. Note, however, that the procedure may not be well suited to very small samples.

[13]Refer to F. White and M. Greig [62], pp. 54-55.

[14]Such a measure would be $R^2 = 1 - S(\hat{B})/S(\bar{B})$, where $S(\hat{B})$ equals the sum of squared residuals with the estimated coefficients, and $S(\bar{B})$ equals the same with the coefficients equal to 0.

[15]McFadden [29] references Theil, "On the Estimation of Relationships Involving Qualitative Variables," *American Journal of Sociology* 76, pp. 103-154, for proof of the procedure.

[16]Since the logit is equal to the index $G(x)$, $\ell n\, P_j^a/(1-P_j^a) = G(x)$, this can also be taken to be the mean of the distribution of the logits across each sample.

[17]i.e., walk times. These results could indicate that individuals living farther away from work park close to the office and pay a higher parking charge (perhaps to ensure the availability of space), while those who live closer trade off walking and parking.

[18]The diagrammatic exposition is borrowed from P. Watson and R. Westin [59].

[19]Westin [59] calculates an error allowance of ±1.96 standard deviations as an appropriate error allowance.

[20]Refer to Chapter 3 and footnote 4 of that chapter.

[21]For example, a reclassification of approximately 4% of the population would have equalized the city-wide elasticity and the aggregate elasticity.

Chapter 6

SUMMARY AND CONCLUSIONS

6.1 SUMMARY OF MAJOR OBSERVATIONS

Several observations have been made in this thesis about the use of disaggregate models, appropriate aggregation procedures, and the transferability issue. The major observations are summarized below.

1. Disaggregate models perform well in identifying aspects of observed behavior, and in calculating individual probabilities of choosing a particular action.
2. Aggregation procedures which take into account the distributions of individual probabilities in a sample differ in their predictions from those generated through naive aggregation procedures. This results in more accurate modal split estimates.
3. Modal splits and calculated elasticities based upon an actual known distribution of probabilities for a sample are similar to those generated (with decreased information requirements) through an aggregation procedure developed by Westin [60].
4. Elasticity measures derived through appropriate aggregation procedures borrow parameters of observed preferences and of distributions of characteristics. This provides sensitivity to the elasticity of different modeling situations. These elasticities, for this reason, will vary across the urban space, at least because of differences in characteristics distributions. This means that city-wide policy might elicit different responses (for city-wide

policy) from different areas. Possibly those elasticities might also vary because of a diversity of measured parameters of observed preferences across spatially-differentiated models.

5. Theoretically, the coefficients of the model (the parameters of observed preferences) are equal. In practical terms, the measured coefficients of the model should be stable across diverse groups of individuals as long as the model is properly identifying only observed preferences. This will hold true if there is no violation of the Axiom of the Independence of Irrelevant Alternatives.
6. If transferability is confirmed, then information can be borrowed from different models as they become available in procedures to make predictions about different populations.
7. If transferability is not expected, then either different models will have to be estimated for the populations in question, or else the coefficients that one has available will have to be updated to suit the new population.

The results of the empirical analysis of Chapter 5 offer considerable support for these observations.

1. All nine models estimated show reasonable performance (as indicated by the results of the Likelihood Ratio Tests which indicate equation significance in all cases of greater than 97.5% confidence). All policy-related variables are of reasonable sign and significance (except those for pcost where measurement difficulties, as expected, reduced the significance of the parameters). The pseudo-R^2 for each model indicate reasonable fit between the predicted and observed individual actions, as the measures range from 0.1 to 0.4, which McFadden [29] considers to be good model performance.

2. The estimated proportion of auto users calculated from naive aggregation procedures overstates the expected proportion of auto users estimated from the model, and also the observed proportion of auto users. The naive aggregation technique using the average values of explanatory variables predicts an expected proportion of auto users equal to 0.6497, compared to a predicted *and* actual proportion of 0.60656. Elasticities calculated without inspecting the distribution of probabilities in the population overstate the population sensitivity to these variables. A comparison of city-wide elasticities calculated under both methods confirms these expectations.

TABLE 6.1
NAIVE VERSUS CORRECTLY CALCULATED CITY ELASTICITIES

ELASTICITIES	MIN(T_A-T_B)	COST(C_A-C_B)	PCOST
Previously calculated model elasticities	0.4777	-0.0358	-0.2848
naive $E = (1-\bar{P}_j)\beta_{jk}(x_{jk})$	0.5296	0.0397	0.3156

The improperly calculated elasticity measure overstates the more accurate elasticity measure by approximately 10.8% in each case.

3. The results of Table 5.24 confirm the validity of Westin's aggregation procedure. For instance, the expected properties of auto users from the actual probability distribution is similar to that predicted from the generated probability distribution for that area.
4. The elasticities calculated from each individual model vary as one moves concentrically away from the city centre. This indicates differential sensitivity to policy variables, which is dependent upon location. The elasticities with respect to time and cost

generally increase as one moves away from the city centre--indicating higher absolute time differences and parking charges facing those who live farther out from the city centre.

When one holds constant all the components of the city-wide elasticities except for parameters of the characteristics distributions, which are allowed to vary for each area submodel, the pattern of elasticities is replicated, although that variation is reduced in magnitude. This indicates that these elasticities vary at least because of the difference in the choice situation facing individuals in different parts of the city. (The rest of the differences in variation between the elasticities is due to holding constant the mean and variance of the probability distribution between areas. These measures will differ across the city even if the coefficients are equal, so this artificially reduces the observed variation in elasticities in such a simple example.)

5. Tests of transferability (stability of the coefficients) between areas indicate that generally the idea of transferability cannot be discounted, and also that the extreme case of nontransferability can easily be explained as a violation of stated transferability preconditions. The coefficient of time differences between auto and bus in areas 3 (6.9 - 10.8 mi) and 3^* (8.1 - 12 mi) is clearly different than the rest of the city. In this case, the variability between auto and bus time indicates that there is a possibility of supply competition on the part of public officials to reduce public transit disadvantage. This competition is either in terms of truly unobservable variables, or else in terms of potentially observable variables such as comfort and convenience which could not be

captured in this model.

Important conclusions are, then, that time and cost differentials could hold significant implications for expected transferability, and that the most successful results would be predicted in areas where either city size or planning means that most determinants of service are in time and cost savings. More importantly, the inclusion of LOS variables other than time and cost will reduce the specification error above, and increase the likelihood of model transferability. This indicates that one research and modeling priority must be in terms of capturing qualitative LOS effects in models.

Comparison of actual elasticities with those generated by using borrowed coefficients, and also moments of a generated distribution of probabilities using these moments (Table 5.), indicates that these are similar. This supports the concept of model transferability, and that the elasticities vary across urban space because of changes in characteristics levels. This also lends support to the aggregation procedure of Westin.

6.2 IMPLICATIONS FOR THE POLICY MAKER

Two potential prediction requirements could exist for a policy maker:

1. A need for policy predictions for city subareas or subpopulations to examine differential urban response to policy.
2. A need for accurate city-wide predictions for the overall results of city policy.

6.2.1 Making Predictions for City Subareas

Given the general success of the preceding tests of

transferability, the following modeling options are open to the policy maker.

1. Estimate a separate model for each subarea being examined.
2. Update the coefficients of a base model to more correctly reflect the parameters expected in the subarea. It would be possible to use these to generate an expected probability distribution, for use in the calculation of elasticities and other policy-related indicators.
3. Assume transferability of coefficients from a base model, and by using these parameters, generate a distribution of probabilities for the area as in 2 above for use in elasticity calculations.
4. Apply simple correction procedure to elasticities generated by a base model to reflect their difference between populations due to changes in characteristics distributions.
5. Use the city-wide model for each subarea.

These options have been listed in decreasing order of the precision of their estimates (measured in terms of their ability to replicate the actual subarea behavior). However, an important consideration in encouraging the adoption of disaggregate procedures is the cost of the procedure and the sample required. These options have also been listed in decreasing order of their data requirements. This emphasizes once again the important trade-off between precision and practicality.

The estimation of a new model for each area necessitates a reasonably sized data set for each area. The requirement of generating such information has been the cause of slower adoption of these procedures in the first place, relative to what one would hope, given their advantages over traditional procedures. This, however, is only

a relevant problem for policy makers who desire to make predictions for areas not covered by a base sample. If the data are available, as has been true in this analysis, this procedure should be considered the best alternative as it is, of course, possible to estimate the model on a subset of existing data.

Updating procedures require at least a small sample for the area; but it does provide a method of correction to suspect model results. This would be especially useful for correcting the parameter estimates for the time variable for areas 3 and 3^*. However, if one wishes to make aggregate predictions with these coefficients, it will be necessary to generate Westin's S_B distribution of probabilities. This, along with the need for at least a small data set, can reduce the attractiveness of this procedure.

The use of base model coefficients alone in Westin's procedure needs significantly less information if transferability is reasonably assumed. One needs only the coefficients of the base model and summary distributional information about the characteristics. This procedure importantly does not need a new data set for the new population. However, the procedure necessitates the programming ability to generate the distribution, which can be costly. However, results from Table 5.27 show that this is a relatively accurate procedure, and needs only the base model results borrowed under reasonable assumptions of transferability. The program generation of this distribution of probabilities for this analysis did prove to be a rather complex and time consuming procedure.

If such procedures are not available, or if cost or time constraints make the procedure prohibitive, then the policy maker can account

for different elasticities for different modeled areas by at least correcting base model elasticities for heterogeneity in the population. The structure of the elasticity allows for this correction, and the analyst needs only the means of the distribution of characteristics, which should be readily available. It is noted that this procedure uses the same information as in other preceding options, but avoids the generation of a new distribution of probabilities. Since this procedure imposes assumed equality of moments of the actual probability distribution, this simple correction is most legitimately done between areas in which the properties of individuals choosing a particular alternative is similar. To apply the correction to areas where there is a higher proportion of individuals using the auto would lead to the calculation of elasticities much overstating the real elasticity. This would be especially true if the variance of the actual probability distribution is low, such as might be the case with areas quite far from the city centre. To correct the elasticity in this manner where there is a much larger proportion of bus users would understate the true elasticity, and this would be especially true if there was a high variance in the individual probabilities.

The easiest, but least accurate, methodology would be to use the city-wide predictions for all areas. This compromises the individuality of different subareas or subpopulations, at least because the city-wide distribution of characteristics is not equal to the subpopulation distributions.

Given the options above, even the most simple correction requires little information, as summary information about characteristics distributions should be available. However, if existing data inventories

include only city-wide aggregate data, the summary statistics for different subpopulations will be difficult to find, and makes even a small amount of correction difficult to apply.

6.2.2 Making City-Wide Predictions

The use of a city-wide model enbodies the assumption that differences in subarea responses does not invalidate the predictions of a city-wide model. Different areas, as indicated in Chapter 5, generate different responses to even city-wide policy. Is the city-wide model accurate in capturing the overall response?

One method of examining this question is to compare the city-wide elasticities generated from the "city" model, to calculated aggregate elasticities indicated by McFadden [21] and Domencich and McFadden [9]. The aggregate elasticity as calculated by McFadden is equal to the sum of the component elasticities, weighted by the share of the "component" in the overall population.

Table 5.30 shows that the cost elasticities from the two methods are almost exactly the same, but that differences exist for parking and time. The parking elasticities are of questionable significance, given that the elasticities are themselves based upon insignificant parameters, and therefore, being random variables also, the elasticities may be insignificant. Measurement problems also reducing the validity of the parking results have also been previously mentioned. The differences between the time elasticities show 1% overstatement for 100% changes in time difference values- This is small error despite the clear nontransferability of the time coefficients for areas 3 and 3^*. One suspects then that the city-wide model performs rather well,

especially given some of the difficulties of correcting for the problem. This leads again to the increased importance of model specification in variables such as qualitative LOS, which have, for instance in this case, been left out of the model specification.

Had the elasticity in area 3, for instance, been approximately 0.5913 instead of 0.6188, the two elasticities would have been identical, and the city-wide model would have captured accurately the total effect of the responses from all areas. It is possible that improved model specifications would have resulted in a lower elasticity for the time variable, thus achieving this result.

In summary, policy makers are forced to trade off sophisticated correction techniques with their costs and with the costs of modeling errors from simpler techniques.

Model transferability provides a reasonable simplification to practical problems associated with the adoption of disaggregate techniques for predictions in transportation. However, the application of techniques embodying transferred results is dependent upon reasonable grounds that this is valid in the modeling issue at hand. This necessitates, firstly, a reasonable theoretical understanding of disaggregate models on the part of potential model users. Secondly, this necessitates further research into methods of increasing the generality and flexibility of disaggregate procedures without violating preconditions for transferability. Lastly, and perhaps most importantly, research (*which is accessible by policy makers*) is needed to determine the sensitivity of aggregate prediction to changes in coefficient values, which will indicate to policy makers some measure of marginal benefit of pursuing sophisticated and perhaps costly modeling procedures

or correction procedures to ensure modeling transferability. This topic has not be pursued in this thesis, but there is some information already available on this topic [18] [61], and it has been recognized [18] as one of the highest priorities facing transferability research.

BIBLIOGRAPHY

1. Algers, S., Hanson, S. and Tegnér, G. "Role of Waiting Time, Comfort and Convenience in Modal Choice for Worktrips." *Transportation Research Record*, #534, 1975, pp. 38-51.
2. Atherton, T. and Ben Akiva, M. "Transferability and Updating of Disaggregate Travel Demand Models." *Transportation Research Record*, #610, 1976, pp. 12-19.
3. Baumol, W. J. *Economic Theory and Operations Analysis.* 4th Edition. Prentice-Hall Inc., New Jersey, 1977.
4. Berkson, J. "Application of the Logistic Function to Bioassay." *Journal of the American Statistical Association*, 39, pp. 357-365.
5. Charles River Associates Inc. *A Disaggregated Behavioral Model of Urban Travel Demand.* Federal Highway Administration, Washington, 1972.
6. ________. *Disaggregate Travel Demand Models, Project 8-13, Phase I Report*, Vol. I, Feb. 1976.
7. ________. *Disaggregate Travel Demand Models, Project 8-13, Phase I Report*, Vol. II, Feb. 1976.
8. De Leuw Cather Ltd. *The Impact of Increased Parking Charges Within the Ottawa/Hull Central Area.* Working paper for the Transportation Development Agency, June 1976.
9. Domencich, T. and McFadden, D. *Urban Travel Demand - A Behavioral Analysis.* A Charles River Associates Research Study, North

Holland Publishing Co., 1975.

10. Flaherty, C. and Mangan, D. "Bus Passenger Waiting Time in Central Areas." *Traffic, Engineering and Control*, Vol. II, No. 9, 1970.

11. Gillen, D. W. "Alternative Policy Variables to Influence Urban Travel Demand." *Canadian Journal of Economics*, No. 4, Nov. 1977, pp. 686-695.

12. ________. *Effects of Changes in Parking Prices and Parking Restrictions on Urban Transport Demands and Congestion Levels.* Research Report #25, University of Toronto, York University Joint Program in Transportation, Feb. 1975.

13. ________. "Effects of Parking Costs on Urban Transportation Modal Choice." *Transportation Research Record*, #649, 1977, pp. 46-51.

14. ________. "Estimation and Specification of the Effects of Parking Costs on Urban Mode Choice." *Journal of Urban Economics* 4 (2) 1977, pp. 186-199.

15. Goldberger, A. "Correlations Between Binary Outcomes and Probabilistic Predictions." *Journal of the American Statistical Association* (68), 1973, p. 84.

16. Heckman, J. J. "Introduction." *Annals of Economic and Social Measurement*, Vol. 5, No. 4, Fall 1976, pp. 1-3.

17. Hensher, D. A. 'Market Segmentation as a Mechanism in Allowing for Variability in Travel Behavior." *Transportation* (5) 1976, pp.

18. Hensher, D. and Stopher, P. "Behavioral Travel Modelling." *Behavioral Travel Modelling* (Hensher and Stopher, ed.), Croom Helm Ltd., London 1979, pp. 11-54.

19. Johnson, N. L. "Systems of Frequency Curves Generated by Methods of Translation." *Biometrika*, Vol. 36, 1949, pp. 149-176.

20. Khalil, D.A.F. "Transportation System Level of Service Indices - A Behavioral Approach," M.Sc. Thesis, Department of Civil Engineering, University of Calgary, 1978.

21. Kirshner, D. and Talvitie, A. "Specification, Transferability and the Effects of Data Outliers in Modelling the Choice of Mode in Urban Travel." *Transportation* (7) 1978, pp. 311-332.

22. Koppelman, F. S. "Guidelines for Aggregate Travel Prediction Using Disaggregate Choice Models." *Transportation Research Record*, #610, 1976, pp. 19-24.

23. ________. "Predictions with Disaggregate Models - The Aggregation Issue." *Transportation Research Record*, #527, 1974, pp. 73-80.

24. Kraft, G. "Economic Aspects of Urban Passenger Transportation." *Highway Research Record*, #285, pp. 10-19.

25. Lancaster, K. "A New Approach to Consumer Theory." In Guendt, *The Demand for Travel, Theory and Measurement*. D. C. Heath and Co., Lexington, Mass., 1970.

26. Luce, R. D. *Individual Choice Behavior - A Theoretical Analysis*. John Wiley and Sons, Inc., London, 1959.

27. Luce, R. D. and Raiffa, H. *Games and Decisions, Introduction and Critical Survey*. John Wiley and Sons, Inc., New York, 1967.

28. McFadden, D. "A Comment on Discriminant Analysis Versus Logit Analysis." *Annals of Economic and Social Measurement*, Vol. 5, No. 4, Fall 1976, pp. 551-524.

29. ________. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers of Econometrics* (P. Zarembka ed.).

Academic Press, 1974, pp. 105-139.

30. ________. "The Measurement of Urban Travel Demand." *Journal of Public Economics* (3) 1974, pp. 303-328.

31. ________. "Quantal Choice Analysis, A Survey." *Annals of Economic and Social Measurement*, Vol. 5, No. 4, Fall 1976, pp. 363-390.

32. ________. "Quantitative Methods for Analysing Travel Behavior of Individuals--Some Recent Developments." *Behavioral Travel Modelling* (Hensher and Stopler eds.), Croom Helm Ltd., London, 1979, pp. 279-316.

33. McFadden, D. and Reid, F. "Aggregate Travel Demand Forecasting from Disaggregated Behavioral Models." *Transportation Research Record*, #534, 1975, pp. 24-37.

34. Mood, A. M. and Graybill, F. A. *Introduction to the Theory of Statistics*. McGraw-Hill Book Co., New York, 1963.

35. Morrison, D. G. "Upper Bounds for Correlation Between Binary Outcomes and Probability Predictions." *Journal of the American Statistical Association* (67) 1972, pp. 68-70.

36. Nerlove, M. and Presse, J. *Univariate and Multivariate Loglinear and Logistic Models*. Report #R-1306-EDA/NIH, Rand Corporation, California, Dec. 1973.

37. Nicholaidis, G., Wachs, M., and Golob, T. "The Evaluation of Alternative Market Segmentations for Transportation Planning." *Transportation Research Record*, #649, 1977, pp. 23-31.

38. Norshek, M. and Sonnenschein, H. "Marginal Consumers and Neoclassical Demand Theory." *Journal of Political Economy*, Vol. 87, No. 6, Dec. 1979, pp. 1368-1376.

39. Oum, T. "A Warning on the Use of Linear Logit Model in Transport

Mode Choice Studies." *Bell Journal of Economics*, Vol. 10, No. 1, 1979, pp. 374-390.

40. Pearce, W. *A Contribution to Demand Analysis.* Clarendon Press, Oxford, 1964.

41. Pindyck, R. and Rubenfeld, D. *Econometric Modeling and Economic Forecasts.* McGraw Hill Book Co., New York, 1976.

42. Polinsky, A. M. and Shavell, S. "Amenities and Property Values in a Model of an Urban Area." *Journal of Public Economics* (5) 1976, pp. 119-129.

43. Rassam, P., Ellis, R., and Bennet, J. "The n-Dimensional Logit Model - Development and Application." *Highway Research Record*, #369, 1972, pp. 135-147.

44. Reichman, S. "Instrumental and Life Cycle Aspects of Urban Travel Behavior." *Transportation Research Record*, #649, 1977, pp. 38-41.

45. Reichman, S. and Stopher, P. "Disaggregate Stochastic Models of Travel Mode Choice." *Highway Research Record*, #369, 1972, pp. 91-104.

46. Rosen, S. "Hedonic Prices and Implicit Markets - Product Differentation in Price Competition." *Journal of Political Economy*, Vol. 82, No. 1, Feb. 1974, pp. 34-55.

47. Stopher, P. "Development of Market Segments of Destination Choice." *Transportation Research Record*, #649, 1977, pp. 14-22.

48. Stopher, P. Spear, B., and Sucher, P. "Towards the Development of Convenience Measures for Transportation Models." *Transportation Research Record*, #527, pp. 16-32.

49. Stopher, P. and Lisco, P. "Modelling Travel Demand, A Disaggregate

Behavioral Approach - Issues and Applications." *Transportation Research Forum*, 1970, pp. 195-214.

50. Stopher, P. and Meyburg, A. *Urban Transportation Modelling and Planning*. D. C. Heath and Co., Lexington, Mass., 1975.

51. Talvitie, A. "Comparison of Probabilistic Mode Choice Models, Estimation Methods and Systems Inputs." *Highway Research Record*, #392, 1972, pp. 111-120.

52. Tardiff, A. "Comparisons of Effectiveness of Various Measures of Socioeconomic Status in Models of Transportation Behavior." *Transportation Research Record*, #534, 1975, pp. 1-9.

53. Theil, H. *Principles of Econometrics*. John Wiley and Sons, Inc., New York, 1971.

54. Travis, K. E. "The Sensitivity of Parameter Estimates to Data Specification in Mode Choice Models." *Transportation* (7) 1978, pp. 301-309.

55. Warner, S. *Stochastic Choice of Mode in Urban Travel - A Study in Binary Choice*. Northwestern University Press, 1962.

56. Watson, P. L. "Predictions of Intercity Mode Choice from Disaggregate Behavioral Stochastic Models." *Highway Research Record*, #446, 1973, pp. 28-37.

57. ________. "Problems Associated with Time and Cost Data used in Travel Choice Modelling and Value of Time." *Highway Research Record*, #369, 1971, pp. 148-158.

58. ________. *The Value of Time: Behavioral Models of Mode Choice*. D. C. Heath and Co., Lexington, Mass., 1974.

59. Watson, P. L. and Westin, R. B. "Transferability of Disaggregate Mode Choice Models." *Regional Science and Urban Economics* (5)

1975, pp. 227-249.

60. Westin, R. B. "Predictions from Binary Choice Models." *Journal of Econometrics* (2) 1971, pp. 1-16.

61. Westin, R. B. and Manski, C. F. "Theoretical and Conceptual Developments in Demand Modeling." *Behavioral Travel Modeling* (Hensher and Stopher eds.), Croom Helm Ltd., London, 1979, pp. 378-392.

62. White, K. and Greig, M. *Shazam, An Econometrics Computer Program.* Computing Centre, University of British Columbia, Sept. 1979.